

# Do partially disabled people respond to financial incentives to work?<sup>\*</sup>

Tunga Kantarci<sup>†</sup>, Wietse Mesman<sup>‡</sup> and Jan-Maarten van Sonsbeek<sup>§</sup>

24 March 2025

## Abstract

We analyze the effects of financial incentives to resume work built into the Dutch DI scheme for partially disabled individuals who have remaining work capacities. Exploiting the sharp discontinuous reduction of benefits after receiving benefits up to 24 months, we show that financial incentives strongly increase work resumption mostly among high-wage earners for who the incentives are larger, but they also increase claims of benefits from alternative social support programs mostly among low-wage earners. The responses are sensitive to the business cycle but not to the duration of benefit receipt. They are stronger among workers with temporary work contracts and the unemployed compared to the workers with permanent work contracts. These effects are persistent in the long run.

## 1 Introduction

In 2006 the Netherlands reformed its DI system which led to a strong decrease in annual inflow into DI (Van Sonsbeek and Gradus, 2013; Kantarci et al., 2023). However, inflow into the scheme for partially and temporarily fully disabled individuals increased from 17,300 in 2006 to 48,000 in 2023. With this increase, the stock of partially disabled individuals claiming benefits has reached 253,000 in 2023. There are several reasons for this increase, of which the most important are the increasing number of older workers and workers without a permanent contract, who both have a higher DI risk in the insured population. It is interesting to analyze partially disabled individuals because they have remaining work capacity to utilize and may be more likely to engage in moral hazard, especially because it is difficult to reassess work capacity

---

<sup>\*</sup>This research is supported by Netspar under grant number LMVP 2014.03. Its contents are the sole responsibility of the authors. We thank the Employee Insurance Agency (UWV), and in particular Lucien Rondagh, Willy van den Berk, Carla van Deursen, and Roel Ydema, for providing the disability data. We thank Hans Bloemen, Ziwei Rao, Arthur van Soest and the participants of the Netspar Pension Day 2021, the Netspar International Pension Workshop 2023, the annual conferences of the Swiss Society of Health Economics 2023 and the European Society for Population Economics 2024 for their helpful comments and suggestions. Results are based on calculations by the authors using non-public microdata from Statistics Netherlands. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information: [microdata@cbs.nl](mailto:microdata@cbs.nl).

<sup>†</sup>Department of Economics, Econometrics and Finance, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands, and Netspar, (e-mail: t.kantarci@rug.nl)

<sup>‡</sup>Department of Economics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, and Netspar (e-mail: w.r.b.mesman@tilburguniversity.edu)

<sup>§</sup>Department of Budget and Taxation, Netherlands Bureau for Economic Policy Analysis, P.O. Box 80510, 2508 GM The Hague, The Netherlands, and Netspar (e-mail: j.m.van.sonsbeek@cpb.nl)

of large numbers of DI recipients. Policy measures could be most effective if they target this group.

The size of a DI scheme can be reduced by either limiting inflow into the scheme or increasing outflow from it. Inflow can be limited by stricter screening during the sickness insurance period, or stricter eligibility criteria for DI at the end of it so that fewer insured workers qualify for DI. Outflow can be increased by stronger reintegration measures. Financial incentives are another measure commonly used in DI schemes to limit inflow or encourage outflow.

A substantial body of literature has developed analyzing the causal effects of financial incentives which can be divided into two strands. The first strand studies the impact of financial incentives on limiting inflow into DI. To limit inflow into DI, DI schemes typically reduce benefit generosity to make DI unattractive to apply. Studies conducted in Austria, Canada and the United States find that reducing benefit generosity reduces benefit claims and awards and increases labor participation, but the sizes of the estimated effects vary considerably ([Bound and Burkhauser, 1999](#); [Gruber, 2000](#); [Mullen and Staubli, 2016](#); [Favre et al., 2021](#)). [Campolieti \(2004\)](#), however, finds no effect on DI claims or labor participation among older workers.

The second strand studies the impact of financial incentives on increasing outflow from DI. To increase outflow from DI, many DI schemes allow earnings up to a certain threshold and reduce benefits if earnings exceed the threshold. Therefore, the studies that focus on outflow investigate the effects of changing the earnings threshold or the effects of reducing benefits at different rates when earnings exceed the threshold. With respect to changing the earnings threshold, in Canada, [Campolieti and Riddell \(2012\)](#) study the effect of changing the amount of wages that DI beneficiaries are allowed to earn without losing DI benefits. They find a substantial labor participation increase of 5 to 6 percentage points (pp) for men and 8 to 10 pp for women due to the introduction of the earnings allowance. [Vall Castelló \(2017\)](#) analyze a reform in Spain on income tax exemption for partially disabled beneficiaries who did not work. The reform eliminated the tax exemption for those aged 55 or younger but kept it for older workers. The reform increased labor participation of the younger group by 6.5 pp. In Austria, if earnings exceed a threshold amount, DI benefits are reduced by up to 50%. [Ruh and Staubli \(2019\)](#) find substantial and sharp bunching at the earnings threshold, and that the average DI recipient with earnings just below the threshold would increase earnings by 45% if the threshold was eliminated. In Canada, [Zaresani \(2020\)](#) exploits a policy reform that changed the earnings threshold to increase labor participation. [Zaresani](#) also finds substantial bunching, which remained persistent after changing the earnings threshold, suggesting substantial adjustment costs for disabled workers.

With respect to reducing benefits when earnings exceed a threshold amount, in the US, [Weathers and Hemmeter \(2011\)](#) study a pilot project that replaced the complete loss of DI benefits due to earning above a threshold amount with a gradual reduction of DI benefits by 50% of the amount earned above an earnings disregard level. They find that the number of DI recipients with earnings above the threshold amount increases by 25% after the policy change. [Kostøl and Mogstad \(2014\)](#) analyze the consequences of providing financial incentives to DI recipients to stimulate work resumption. In 2005, the Norwegian government introduced a program where DI benefits are reduced by 60% of the earnings accumulated above an earnings threshold. They find a substantial positive effect on labor participation.

Our study focuses on the second strand of the literature analyzing the impact of financial incentives to increase outflow in the existing stock of DI recipients. In the Netherlands, when admitted to the DI scheme, partially disabled individuals start receiving a DI benefit that incorporates in it the unemployment insurance (UI) benefit. The benefit is reduced at a rate of 70% if benefit recipients earn wages. The duration of the benefit depends on the work history and expires after 3 to up to 24 months from the first day of DI receipt. After this period

individuals face two financial incentives to increase labor supply. First, the UI component of the benefit expires and DI recipients can no longer claim UI. Second, if their earnings exceed an individual-specific threshold amount, their benefit is increased so that an increase in earnings leads to an increase in the benefit. The financial incentives are larger for benefit recipients who earn higher wages before they fall sick. To compensate for lost benefits, individuals may increase labor supply although they may fall back on alternative social support programs.

Koning and van Sonsbeek (2017) evaluate the impact of these financial incentives among partially disabled individuals who claimed disability benefits from 2006 to 2013. They find an average labor participation increase of 1.4 and 2.6 pp in the short- and long-term, respectively. They find no effect on full work resumption. These effects are small compared to the international literature but also considering the large size of the financial incentives. We, once again, evaluate the impact of the financial incentives but using extended data from 2006 to 2020. We take a regression discontinuity approach to identify causal effects and show that our estimates are robust to alternative identification strategies.

We contribute to the literature with several findings. First, we show that, for the average wage earner (before falling sick), the financial incentives increase labor participation by 4.5 pp. Netting out the anticipation and adaption effects of the incentives, this estimate increases to about 9 pp. These effects are much larger than those previously found by Koning and van Sonsbeek (2017). We show that this is due to the business cycle that has improved during the extended observation period of our study, after two financial crises between 2007 and 2010. This finding is plausible because disabled individuals have limited job opportunities and can become susceptible when job opportunities become scarce. This suggests that evaluations of reintegration policies should account for the business cycle.

Second, in line with how the financial incentives are designed to be larger for beneficiaries who earn higher pre-sickness wages, we show systematic heterogeneity of the impact of the financial incentives across the (pre-sickness) wage distribution. Considering the two ends of the wage distribution, those who earn at most the minimum wage increase labor participation by about 2 pp while those who earn at least 2.5 times of the minimum wage increase labor participation by about 5 pp after facing the financial incentives for one year. As beneficiaries of different pre-sickness wage groups cannot manipulate the size of the incentives, this shows that the larger the income effect of the financial incentives, the larger the responses. It also shows that the DI scheme that seems to put the financial incentives right ex ante, while serving to its primary objective of income replacement, is still ineffective in inducing work resumption in the vulnerable and large group of low-wage earners who already face limited job opportunities.

Third, several studies showed that reforms that limit inflow into DI through stricter eligibility rules lead sick individuals to substitute DI benefits with earnings but also benefits from alternative social support programs (Karlström et al., 2008; Staubli, 2011; Kantarci et al., 2023). Borghans et al. (2014) show benefit substitution due to reassessment of disability grades of existing DI recipients in the Dutch DI scheme in 1993 based on stricter entrance criteria. They explain that targeting existing DI recipients is of importance for policy interventions because effects which operate on the existing stock of recipients have the potential to make a much greater immediate impact than effects which operate on the comparatively small inflow into a DI scheme. We add to this finding with evidence on the impact of financial incentives among existing DI recipients. Partially disabled individuals strongly respond to benefit reductions by increasing labor participation but also by substituting DI benefits with benefits from alternative social support programs. Those who respond with labor participation and with benefit substitution, however, appear to be fairly segmented groups of high- and low-wage earners, respectively.

Fourth, recent studies compare the effectiveness of inflow and outflow related measures to

limit DI claiming (Van Sonsbeek and Gradus, 2013; Haller et al., 2024; Kantarci et al., 2023). They argue that inflow related measures are more effective. These two types of measures apply at different stages of sickness, however. A particular measure might be effective at the early stages of sickness but may lose its effectiveness as the sickness progresses. Such duration dependence is typically ignored in evaluations or comparisons of policy measures. We provide first evidence that financial incentives to resume work are equally effective for different durations of DI claiming.

Finally, we show that temporary contract workers and the unemployed respond to the financial incentives much more strongly than those with permanent work contracts. This shows that these groups have substantial unused work capacity that gets activated with financial incentives. Burkhauser et al. (2014) recognize that even people with severe impairments can work up to some extent. Our finding contributes to this recognition by highlighting the heterogeneity in work capacity based on contract type.

The remainder of this paper is structured as follows. Section 2 describes the Dutch DI scheme for partially disabled individuals and the financial incentives to resume work. Section 3 describes the data. Section 4 presents descriptive statistics and evidence on the impact of the financial incentives. Section 5 describes the identification strategy. Section 6 presents the estimation results and tests the identifying assumptions. Section 7 concludes.

## 2 The Dutch DI scheme and the financial incentives for work resumption

The Work and Income According to Labor Capacity Act (WIA) came into effect on 1 January 2006 for people who fell ill from 1 January 2004 onwards. In the WIA, individuals who lose any part of their earning capacity due to a health impairment are entitled to the sickness benefit from their employer for a period of two years. When the sickness benefit expires, they can apply for a DI benefit. If the wage loss is more than 80% and expected to be lasting, the individual is admitted to the Income Provision Scheme for Fully Disabled People (IVA). We do not study IVA recipients since their remaining earning capacity is very limited and odds of work resumption are very small. If the wage loss is more than 35% but less than 80%, or if the wage loss is more than 80% but there is a possibility of recovery, the individual is admitted to the Return to Work Scheme for Partially Disabled People (WGA).

The WGA consists of two stages. In the first stage the individual is entitled to the “wage-related benefit”. The benefit is the sum of two components. The first component is a fixed amount (as long as the disability grade is not reassessed) and equal to

$$0.7 \times \text{pre-sickness wage} \times \text{disability grade}.$$

The second component is variable depending on how much the remaining work capacity is utilized and equal to

$$0.7 \times \text{pre-sickness wage} \times (1 - \text{disability grade}) \times (1 - \text{utilization rate}).$$

Utilization rate takes a value between 0 and 1, so that the last term represents the remaining earning capacity that is not utilized.<sup>1</sup> In fact, this component is the UI benefit embedded in the wage-related benefit that compensates individuals who do not utilize their remaining earning capacity.<sup>2</sup> Duration of the wage-related benefit, for this reason, is equal to the duration of the

---

<sup>1</sup>During the first two months of disability, the fraction is 0.75.

<sup>2</sup>It is possible for the utilization rate to be larger than 1 if remaining earning capacity is incorrectly estimated. In this case, the second component of the benefit is negative. If this situation persists, it may lead to a reassessment of the disability grade.

UI benefit. It depends on the number of contribution years to social insurance, with each year awarding half a month of duration. Maximum duration is 24 months.<sup>3</sup> UI benefit entitlement (regardless of disability) is exhausted as the wage-related benefit is exhausted.

When the wage-related benefit expires, individuals become entitled to one of two types of second-stage benefits depending on whether they utilize at least 50% of their remaining earning capacity. If they utilize at least 50% of their remaining earning capacity, they become entitled to the “wage-supplement benefit” which is equal to

$$0.7 \times \text{pre-sickness wage} \times \text{disability grade}.$$

If they utilize less than 50% of their remaining earning capacity, they become entitled to the “follow-up benefit” which is equal to

$$0.7 \times \overline{\text{pre-sickness wage}} \times \text{disability grade},$$

where pre-sickness wage is capped at the minimum wage. Both benefits make flat-rate payments since they disregard how much individuals work once they are below or above the 50% threshold. Both benefits are paid as long as the individual is disabled but expire when the individual becomes entitled to the state pension at the statutory retirement age.<sup>4</sup> The individual can become entitled to UI benefit during the second stage of the WGA, after exhausting their UI benefit during the first-stage. The condition is to have worked 26 weeks in the last year and 4 years in the last 5 years.

These benefit rules imply two financial incentives to increase work effort when the first-stage wage-related benefit expires. The first incentive is due to the reduction in the amount of the DI benefit in the second-stage of the scheme since the UI benefit component of the first-stage wage-related benefit expires and is no longer a component of the second-stage wage-supplement or the follow-up benefits. The second incentive is due to the reduction in the DI benefit from the wage-supplement benefit to the follow-up benefit if individuals do not utilize at least 50% of their remaining work capacity in the second stage of the scheme: the wage-supplement benefit is based on pre-sickness wage while the follow-up benefit is based on  $\overline{\text{pre-sickness wage}}$ . As both incentives reduce the benefit level, they imply a strong negative income effect on labor supply. The size of the benefit reduction in the second-stage of the scheme, however, depends on how much individuals work in the first- and second stages, which implies a substitution effect.

Figure 1 illustrates the two financial incentives for a hypothetical individual who earns €100 per day before reporting sick and has a disability grade of 50%. It shows how the wage, DI benefit, and the sum of the two change with the amount of work. Amount of work is stated in terms of the daily wage during disability as a fraction of the pre-sickness daily wage. During disability, a daily wage equal to 25% of the pre-sickness daily wage implies a remaining work capacity utilization rate of 50% (given the assumed 50% disability grade). The figure distinguishes, with the vertical reference line, between two cases. To the left of the line, the individual utilizes less than 50% of her remaining work capacity in the second stage of the DI scheme and qualifies for the follow-up benefit. To the right of the line, she utilizes more than 50% of her remaining work capacity and qualifies for the wage-supplement benefit. The figure compares the total income from wage income (dashed or solid line in blue color) and a second-stage benefit (dashed or solid line in orange color) to the total income from earnings and

---

<sup>3</sup>The maximum duration has been reduced from 60 to 38 in January 2008, and it has been gradually reduced from 38 to 24 months in 14 quarters between January 2016 and April 2019.

<sup>4</sup>Individuals receive decision letters on DI eligibility from the Employee Insurance Agency on two occasions. First, they receive a letter on their application for the wage-related benefit at the end of their participation in the SI scheme. Second, they receive a letter approximately two months before their wage-related benefit expires which informs on the entitlement decision of a second-stage benefit.

the first-stage wage-related benefit (dashed or solid line in red color) at given remaining work capacity utilization rates.

Two notable patterns in Figure 1 demonstrate the financial incentives. First, a second-stage benefit is always smaller than the first-stage benefit. This demonstrates the first financial incentive: UI benefit is not part of a second-stage benefit. The difference (that is, the incentive) is smaller the more the individual works, and it vanishes if the individual utilizes her remaining work capacity to the full extent in which case the UI component of the first-stage benefit is 0 (where red line intersects orange line).

Second, a second-stage DI benefit is smaller if the individual utilizes less than 50% of her remaining work capacity (and qualifies for the follow-up benefit) than if she utilizes at least 50% of her remaining work capacity (and qualifies for the wage-supplement benefit). This demonstrates the second financial incentive: working less than the 50% threshold remaining work capacity utilization rate reduces the benefit. The vertical distance between the total income from earnings and the first-stage benefit (dashed or solid black line) and the total income from earnings and a second-stage benefit (dashed or solid gray line) shows the extent of the financial incentives to work in the second stage (when the individual utilizes less than or at least 50% of her remaining work capacity).

Individuals who earn higher pre-sickness wages face larger incentives to work because their benefit decreases by a larger amount from the first to the second stage of the DI scheme, implying larger income effects. This is illustrated by comparing Figure 1 to Figure A1 in the appendix. The financial incentives for individuals with high pre-sickness wages are not only larger than the incentives for those with low pre-sickness wages in absolute terms, but also relative to pre-sickness wages. Although the loss of the UI component of the wage-related benefit is proportional to the pre-sickness wage, the penalty for not meeting the threshold remaining earning capacity utilization rate of 50% is based on the difference between the pre-sickness wage and pre-sickness wage capped by the minimum wage (pre-sickness wage). For pre-sickness wages below the minimum wage, this difference is non-existent. Above that, the difference increases with the pre-sickness wage. For these reasons, we conduct all our analyses by distinguishing across pre-sickness wage groups.

The magnitude of the financial incentives may depend on whether the individual is entitled to the “social minimum supplement”. This benefit supplements a social security benefit up to the subsistence-level “social minimum” set by the government. However, the total of the social security benefit and the supplement cannot exceed the pre-sickness wage. If this prevents the individual from reaching the social minimum, she may apply for “social assistance” to supplement the income up to the social minimum.<sup>5</sup> For some individuals the follow-up benefit can be very low so that they may qualify for the social minimum supplement and social assistance, which can compensate part of the reduction in the DI benefit in the second stage of the DI scheme. *Ceteris paribus*, this case is more likely to apply to low pre-sickness wage earners, reducing their smaller incentives to resume work even more.

The financial incentives of the Dutch DI scheme are unique. In Norway and the US, for example, DI schemes allow earnings up to a given threshold, and decrease benefits if earnings exceed the threshold, at a rate of 60% in Norway and 50% in the US. In the first-stage of the Dutch DI scheme, there is no earnings allowance. That is, in case of earnings, the disability benefit is decreased at a rate of 70% (red line in Figure 1). This means that, although additional earnings lead to a higher total income from wages and the disability benefit (black line in Figure 1), the disincentive to work in the Netherlands is larger than those in Norway and the US. In the second stage of the Dutch DI scheme, earnings are allowed up to the full earnings capacity of the

---

<sup>5</sup>If an individual has a partner with an income, she may get less or no social minimum supplement or social assistance. Hence, not everyone with very low (own) income qualifies for these benefits.

beneficiary, and as soon as earnings are higher than 50% of the remaining earning capacity, the disability benefit is increased (comparing solid and dashed orange lines in Figure 1). Moreover, the benefit is reduced in the second stage of the scheme since UI is no more a component of the disability benefit as in the first stage of the scheme (comparing red and orange lines in Figure 1). These are strong financial incentives to resume work that are unique to the Dutch DI scheme.

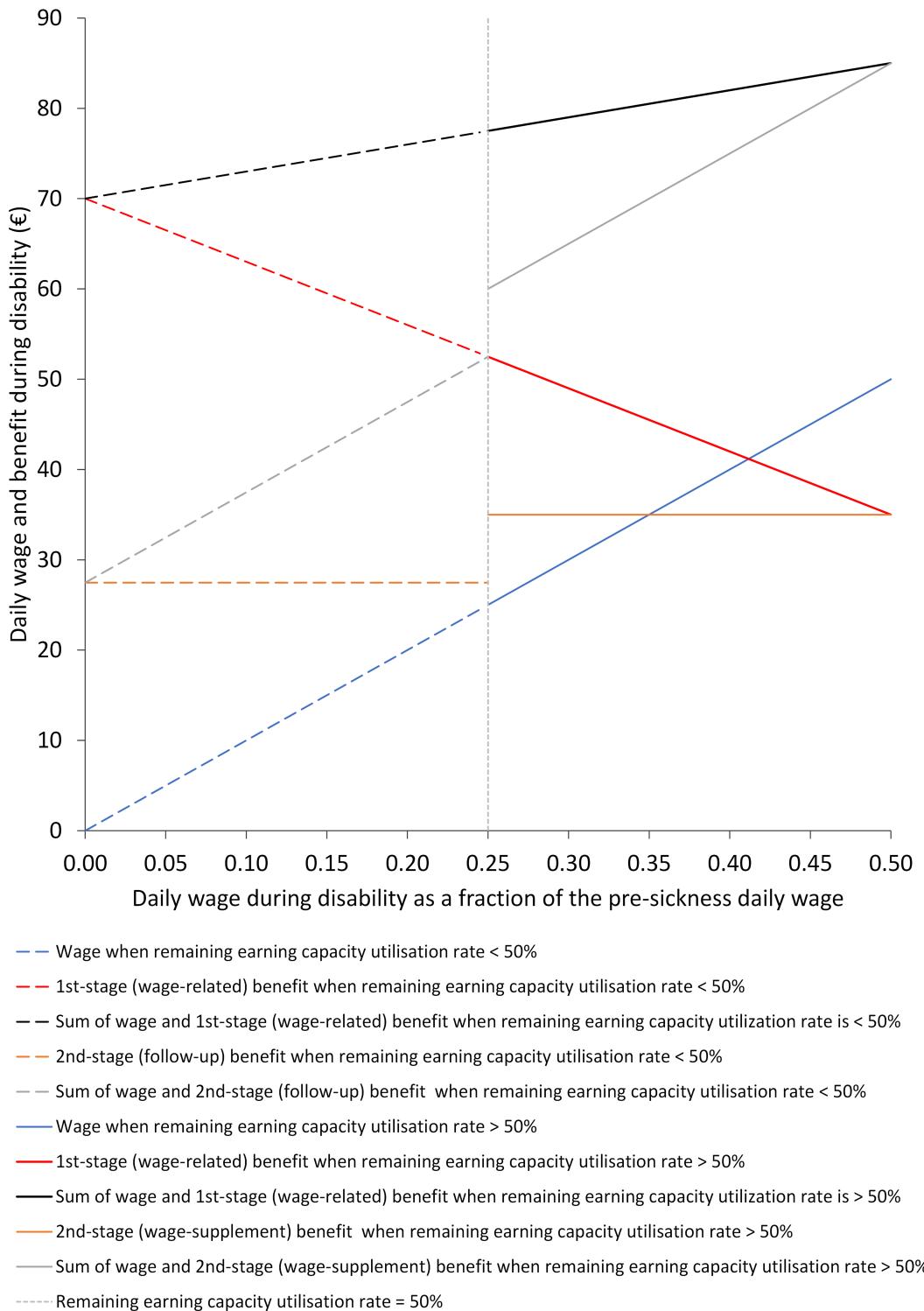


Figure 1: Financial incentives to work in the second stage of the DI scheme for an individual who earns €100 per day before falling sick and has a disability grade of 50%. The financial incentive is the difference between the dashed black line (sum of wage and the wage-related benefit) and dashed gray line (sum of wage and the follow-up benefit) if remaining work capacity utilization rate < 50% in the second stage of the DI scheme (left side of vertical reference line). The financial incentive is the difference between the solid black line (sum of wage and the wage-related benefit) and the solid gray line (sum of wage and the wage-supplement benefit) if remaining work capacity utilization rate  $\geq 50\%$  in the second stage of the DI scheme (right side of vertical reference line).

### 3 Data

We use administrative data from the Employee Insurance Agency on all individuals who participated in the Return to Work Scheme for Partially Disabled People (WGA). The observation period starts in January 2006, when sick individuals could apply for DI benefits the first time in the new DI scheme (WIA), and ends in December 2020. We observe their date of birth, gender, and, on a monthly basis, type of their DI benefit (wage-related, wage-supplement, follow-up), disability grade, and wages earned before reporting sick and during DI receipt. At the time of reporting sick, we also observe whether they received unemployment insurance, and for wage earners, whether they had a permanent contract, a temporary contract, or a contract through a temporary work agency. We link these individuals to administrative data on themselves from Statistics Netherlands, with monthly information on wages and benefits from social support programs from January 2006 to December 2020. Using the wage data, we track their labor participation status after they leave the DI scheme and hence the DI data. Based on these data, we define four outcome variables: dummies for labor participation (based on earning positive wages), social minimum supplement receipt, social assistance receipt, and unemployment insurance receipt.

The initial data set has 17,416,778 observations for 307,356 individuals participating in the WGA. To select the estimation sample, we drop individuals if they are deemed temporarily fully disabled (Section 2). For these individuals remaining earning capacity is 0 meaning that prospects of resuming working, and hence to respond to the financial incentives, is very small among them. More importantly, financial incentives are absent for these individuals. Since their remaining earning capacity is 0, they fulfill the requirement of earning at least 50% of their remaining earning capacity, already entitling them to the higher wage-supplement benefit instead of the lower follow-up benefit in the second stage of the DI scheme. Since they do not receive the unemployment insurance benefit in the first stage of the scheme, they also do not face a reduction in their DI benefit due to the expiration of unemployment insurance in the second stage of scheme. We also drop individuals who re-enter the WGA after exiting. Moreover, we drop those who move back to the first stage of the WGA from the second stage. Finally, we drop individuals if they exit the DI scheme during the first stage of the WGA, or do not reach the second stage by the end of the observation period because these individuals do not experience the financial incentives to resume work. Our study sample has 3,365,133 observations for 46,250 individuals.

### 4 Descriptive statistics

Table 1 presents sample averages of background characteristics for three groups of individuals based on the pre-sickness wages they earned as a fraction of the minimum wage. Individuals with higher pre-sickness wages are older, more often have permanent work contracts, and are more often male. Potential explanations are that older individuals have more job tenure and job security and therefore more often have permanent contracts, and women more often earn part-time wages. In line with these explanations, individuals with higher pre-sickness wages have a longer work history which explains why they receive the first-stage wage-related benefit for a longer period (Section 2). In the second stage of the DI scheme, they also receive the larger wage-supplement benefit for a longer period and the smaller follow-up benefit for a shorter period, most likely because they have better job opportunities and stronger labor market attachment and therefore more often able to utilize more than 50% of their remaining earning capacity to

become entitled to the wage-supplement benefit instead of the follow-up benefit.<sup>6</sup> Individuals with higher pre-sickness wages are more likely to have a higher disability grade but this is because the disability grade is inherently associated with a higher pre-sickness wage.<sup>7</sup>

The top-left panel of Figure 2 shows the fraction of individuals working during two periods of two years each. Months -24 to -1 refer to the period individuals claim the wage-related benefit. Month 0 indicates the first (“cut-off”) month individuals claim the wage-supplement or the follow-up benefit and face the financial incentives to resume work. We distinguish among five pre-sickness wage groups. During the wage-related period, labor participation shows a notable decrease among the two lowest pre-sickness wage groups. It might be that these individuals cannot find suitable jobs to remain active or do not recover from disability to utilize their remaining work capacity. During the wage-supplement or the follow-up benefit period, when the financial incentives become effective, labor participation increases notably for the three highest pre-sickness wage groups. Labor participation of all wage groups show gradual increases in the months just before and after the cut-off month when the financial incentives become effective. These suggest, respectively, anticipation and adaptation effects of the financial incentives.

The top-right and the bottom panels of Figure 2 show the fractions of individuals receiving three types of benefits. For all pre-sickness wage groups, there are obvious and large increases in social minimum supplement receipt at the cut-off month. The social assistance receipt shows small increases at the cut-off month. The lowest wage group is more likely to rely on the social minimum supplement or the social assistance. This is because this group more often becomes entitled to these benefits due to earning low wages or receiving smaller benefits. It is notable that even the highest pre-sickness wage group falls back on these benefits when the DI benefit becomes less generous. Apparently, in all wage groups, some individuals struggle to resume work after they exhaust their wage-related benefit, and rely on benefits from alternative social support programs instead of earnings. It is important to note that enrollment to these benefits is not automatic. Benefit substitution can therefore not occur automatically when DI benefits decrease at the cut-off month.

It is not possible to claim UI while claiming the wage-related benefit (Section 2). Still, a small fraction of the sample claims UI during the wage-related benefit period. It might be that these individuals were already claiming UI when they became partially disabled. UI claims increase during the wage-supplement or follow-up benefit period. This can happen if individuals resume work after they exhaust their wage-related benefit (and hence also their UI benefit), but lose their job after a while and become eligible for UI again.

---

<sup>6</sup>Figure A2 in the appendix shows for the whole sample the distribution of the number of months spent in the first and second stages of the DI scheme.

<sup>7</sup>Disability grade is defined as the difference between the reference wage and assessed remaining earning capacity, divided by the reference wage. The reference wage is the theoretical wage of someone who is not disabled but is similar to the partially disabled individual in all characteristics. The remaining earning capacity is based on a set of jobs that the partially disabled person can still do. For individuals with low pre-sickness wages – and therefore also low reference wages – there will not be jobs paying hourly wages far smaller than their pre-sickness wages since the lowest wage they can earn is the minimum wage. For individuals with higher pre-sickness wages, however, remaining earning capacity can be considerably below their reference wage. Therefore, disability grade can have a strong correlation with pre-sickness wage.

Table 1: Sample averages of background characteristics by pre-sickness wage group

	Pre-sickness wage up to 100% of minimum wage	Pre-sickness wage 150–200% of minimum wage	Pre-sickness wage above 250% of minimum wage
Age	46.163	48.734	53.605
Male	0.203	0.508	0.721
Permanent contract	0.407	0.544	0.554
Temporary contract	0.289	0.214	0.182
Unemployed	0.270	0.238	0.264
Disability grade larger than 50%	0.203	0.510	0.455
Months spent in wage-related benefit	19.607	22.306	25.541
Months spent in wage-supplement benefit	12.803	21.326	22.271
Months spent in follow-up benefit	27.932	22.976	18.814
Pre-sickness daily wage	48.290	102.783	184.370
Number of individuals	2,496	12,318	10,143

Notes: 1. Age, gender, contract type and disability grade are measured in the last month of the wage-related benefit. 2. Pre-sickness daily wage and minimum wage are in January 2006 euros, correcting for indexation of wages with respect to price and sectoral wage inflation.

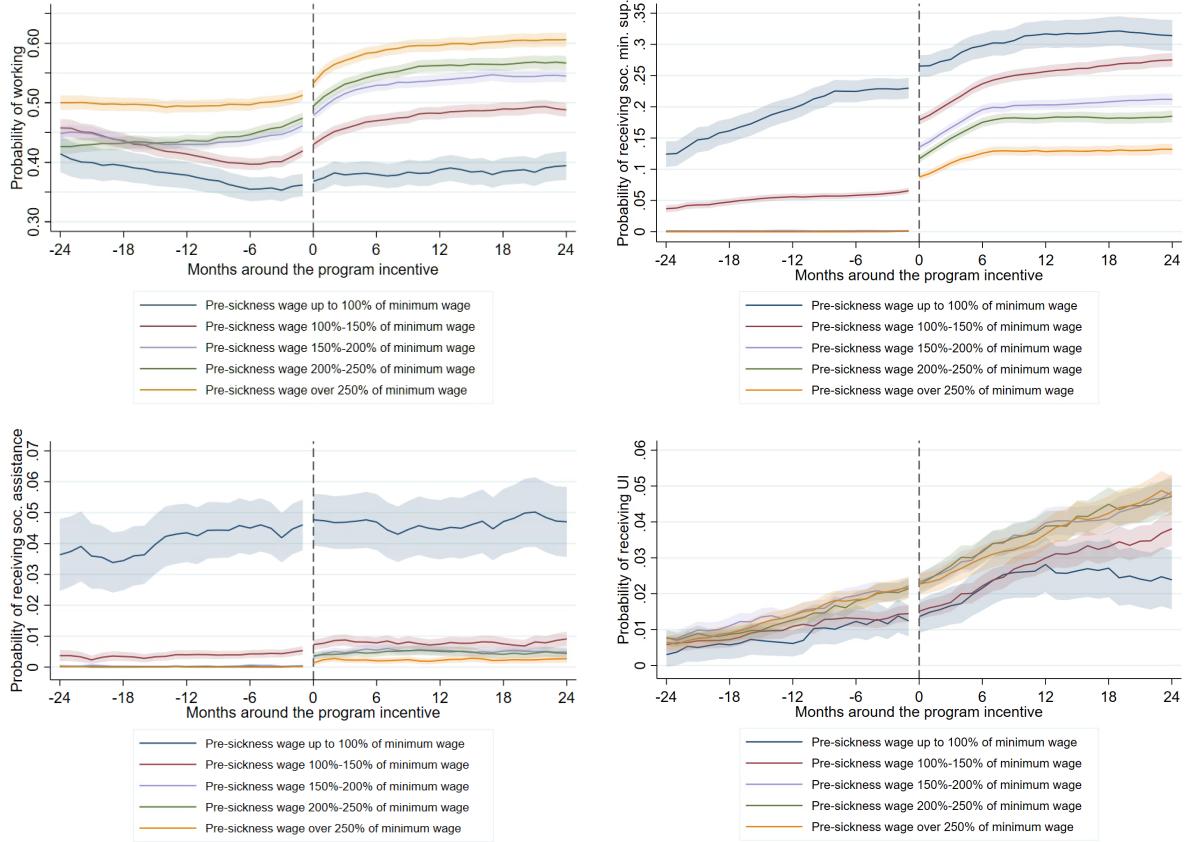


Figure 2: Labor participation, social minimum supplement, social assistance, and UI receipt during the wage-related benefit receipt and after when the financial incentives become effective.

## 5 Identification strategy

As explained in Section 2, individuals with higher pre-sickness wages face a stronger financial incentive to resume work when their first-stage DI benefit expires. Therefore, in baseline but also when studying heterogenous effects of the financial incentives, we distinguish across five pre-sickness wage groups to analyze if the program leads to higher labor participation responses among those with higher pre-sickness wages. We also distinguish across pre-sickness wage groups when analyzing claims of benefits from alternative social support programs.

The estimation strategy to identify the effects of the financial incentives is based on a Regression Discontinuity (RD) design. This design requires a cut-off in the running variable where the probability of treatment changes discontinuously. Our running variable is the time around the moment individuals move from the first to the second stage of the DI scheme. It is defined such that it equals 0 in the first (event) month the individual receives a second-stage benefit (wage-supplement or follow-up benefit), 1 in the month after, -1 in the month before, etc. By this construction, probability of treatment is 0 for all negative values of the running variable, and 1 for all nonnegative values. This gives rise to a sharp RD design as compliance is perfect.

As the running variable takes on discrete values, use of the standard continuity-based approach to RD is questionable, especially given the fact that the number of distinct values of the running variable is as small as 234. That is, when the running variable is discrete and a continuity-based approach is used, the effective number of observations used is the number of the distinct values, not the total number of observations (Cattaneo et al., 2020). For our study, this means estimating the treatment effect using 234 distinct values of the running variable, instead of the available 3,365,133 monthly observations for 46,250 individuals. In this case Cattaneo et al. explain that the local randomization (LR) approach to RD may be the only valid RD approach. This approach requires stronger assumptions than the continuity-based approach. However, if its assumptions are met, it allows the running variable to be discrete and to still identify a causal effect.

The rationale behind the LR identification strategy is that there exists a window around the cut-off where the data behaves as if it were part of an experimental setup in which units receive a random score (value of the running variable), and are assigned to treatment if and only if their score is equal to or above the cut-off value. Consequently, the value of the running variable is not related to any observed or unobserved characteristic of the units, other than by chance. As long as the assigned score by itself does not affect the outcome, the expected value of the outcome variable, conditional on treatment status, is the same for all values of the running variable within the window, even if the control and treatment groups are only observed in (mutually exclusive) parts of the window. This is illustrated in Figure A3 in the appendix. The difference between the expected values of the control and treatment groups identifies the treatment effect. This difference can be estimated by the difference in the sample averages of the two groups.

The LR approach to RD identifies the treatment effect around the cut-off. As we do not take a continuity-based approach to RD, we do not make continuity assumptions to identify and estimate a causal effect at the cut-off (Mattei and Mealli, 2017). Furthermore, Cattaneo et al. (2015), Mattei and Mealli and Sales and Hansen (2020) propose different sets of assumptions within a neighborhood of the cut-off to interpret a RD design as a local randomized experiment. We adapt the approach of Cattaneo et al..

## 5.1 Identification assumptions

Let  $r_0$  be the cut-off value of the running variable,  $T$  the treatment indicator, and  $W_0$  the local randomization window. The potential outcomes of individual  $i$  under treatment status for hypothetical values of the running variable are denoted with  $Y_i(t, r)$ . There are two assumptions that need to hold for the validity of the LR approach:

1. Unconfounded assignment. The distribution function of the running variable inside the window,  $F_{R_i|R_i \in W_0}(r)$ , does not depend on the potential outcomes and is the same for all units:  $F_{R_i|R_i \in W_0}(r) = F_0(r)$ , where  $F_0$  is any distribution function.<sup>8</sup>
2. Exclusion restriction. The potential outcomes do not depend on the value of the running variable inside the window, except via the treatment assignment indicator:  $y_i(t, r) = y_i(t) \forall i$  such that  $R_i \in W_0$ .

The first identifying assumption has two requirements. First, it requires that potential outcomes do not affect the running variable inside the window. This implies that the running variable is not endogenous to potential outcomes. Second, it requires that the distribution of the running variable is the same for all individuals within the window.

Our running variable is determined as follows. For each individual, it starts at a certain negative value (where values denote months) depending on the individual's work history, increases incrementally over (event) time until the individual exploits the first-stage benefit, and, from a value of 0 at the cut-off, it remains increasing in the second-stage until the individual exits the scheme or until it is the end of the observation period.<sup>9</sup>

With respect to the first requirement, we distinguish between the first- and second-stage durations of the DI scheme that span the values of our running variable. Since the duration of the first-stage benefit is institutionally determined, it cannot be adapted in response to the changes in labor supply or benefit claiming, or to factors correlated with these. Therefore, neither the first-stage duration nor the treatment assignment at the end of the first stage can be endogenous. As we control for individual-specific fixed effects, individual attributes that are time invariant and affect the running variable in the second stage through labor supply or benefit claiming cannot confound the treatment effect estimate. For example, individuals with strong relationships with their employers could have better opportunities to resume work when they face the financial incentives of the program in the second stage, and hence have a shorter second-stage duration. However, as long as this is a time invariant attribute, our estimate of the treatment effect cannot be biased.

The second requirement of first assumption is also likely to be met. Individuals cannot manipulate the value of the running variable when they are in the DI scheme. Furthermore, due to the panel nature of the available data, we have observations on (mostly) the same people before and after the cut-off. Therefore, people on both sides of the cut-off should have same or similar (unobserved) characteristics.

There is, however, a caveat regarding the timing of entry into and exit from the DI scheme. If this was the same for all individuals, they would all have experienced same values of the running variable, regardless of their observed (e.g. duration of the first-stage benefit) or unobserved characteristics. In particular, this would imply that their potential outcomes cannot affect the running variable. In fact, this holds for the smallest possible window,  $\{-1, 0\}$ . Everyone in

---

<sup>8</sup>We do not assume that the distribution function is known. This is not needed because in the analysis we do not make use of finite-sample inference due to the large sample available (Cattaneo et al., 2020).

<sup>9</sup>Duration of the first-stage benefit is determined by the duration of unemployment insurance which depends on the work history (Section 2).

the study sample experiences both the last month of the first stage and the first month of the second stage of the DI scheme. However, as the window widens, more and more individuals have missing observations within the window at one or both sides of the cut-off. These missing observations may not be distributed randomly across individuals. The duration of the first stage of the DI scheme, and thereby the number of observations below the cut-off, is determined by the work histories of individuals. Individuals with long work histories may systematically differ from those with short work histories. For example, individuals with higher potential outcomes may have been more likely to work before falling sick and therefore have a longer first-stage duration. Similarly, people who exit the DI scheme soon after the cut-off may be systematically different from those who do not, although a relatively small part of the sample (14.5%) exits the scheme before the end of the observation period, of whom less than half exit in the first 12 months after the cut-off. We expect that such differences are captured by the individual-specific fixed-effects that we control for in our regressions.

The second identifying assumption implies that if there was no treatment, the expected value of the outcome should be the same for all values of the running variable inside the window  $W_0$ . Indeed, due to the exogenous variation in the duration of the first stage, and hence the moment of entry into the DI scheme in terms of calendar time, the month of the cut-off is like any other month, and therefore it is not immediately clear why being in a certain month before or after it should matter for the outcomes (labor participation and benefit claiming) in the absence of treatment.

However, the running variable does contain information about the passage of time. For each individual this variable is perfectly correlated with other variables related to time such as the time spent in the DI scheme, age, and calendar time. These variables may affect the outcomes. Below we discuss how we account for these variables in our regressions.

As mentioned above, exploiting the panel dimension of the available data, we control for individual-specific fixed effects to account for time-invariant observed and unobserved characteristics that may affect the outcomes. Therefore, the two identifying assumptions imply that, conditional on time-variant controls and time-invariant individual attributes, potential outcomes and the running variable are not related except through treatment assignment.

## 5.2 Identification of the full treatment effect

Identification of the full treatment effect regards the timing of responses to the financial incentives and requires that the exclusion restriction assumption holds. That is, since the identification of the treatment effect is based on a comparison of the outcome before and after the cut-off, we need that people close to the cut-off, but still in the first stage of the scheme, do not already change their labor participation or benefit claims in anticipation of the program incentives of the second stage. The modest increase in labor participation of all wage groups just before the cut-off in Figure 2 might cast doubt on this assumption, as it could be explained by anticipation of the incentives.<sup>10</sup>

A similar issue arises for the labor supply response after the cut-off. If individuals are able to utilize their remaining work capacity in suitable jobs at any given point in time, they could react to the program incentives immediately after the cut-off. In practice, however, it may be unrealistic to expect individuals to time their response so precisely. The labor supply response may also be spread out farther beyond the cut-off due to adaptation. In this case the RD estimator will, again, underestimate the true treatment effect.

---

<sup>10</sup>In general, changes in the outcome prior to the treatment could point to endogeneity of the treatment, rather than anticipation (Malani and Reif, 2015). However, as argued above, this is impossible since the duration of the first stage of the DI scheme is institutionally determined.

Underestimation of the treatment effect due to anticipation before the cut-off, or adaptation after the cut-off, is not solved by taking a wider window, although the estimates may be closer to the true effect, because wider windows will still include, albeit to a lesser extent, the anticipation and adaptation effects. We will address this issue in sensitivity analysis by considering a “donut-hole” regression where we will exclude the months in the vicinity of the cut-off month where anticipation and adaptation effects occur.

### 5.3 Window selection

An important step in the LR design is window selection. Ideally, the selected window is the widest possible window within which the LR assumptions hold. However, this cannot be tested directly. [Cattaneo et al. \(2020\)](#) describe two options to select the window. The first option takes a data-driven approach. The window size estimate is the window within which one or more observed covariates do not change with the running variable but can change outside it. Such a covariate does not exist in our data. This leaves us with the second option, which is choosing the window in an ad-hoc manner. We consider a window size that is symmetric around a hypothetical value of  $-0.5$ , and therefore contains the same number of months before and after the cut-off.

Several considerations should be weighed against each other when choosing the window. On the one hand, the smaller the window, the more likely that the LR assumptions hold. Moreover, the smaller the window, the more plausible it becomes that there is no relation between the running variable and the outcome variable anywhere in the window. In the context of the current study, the further away in time from the cut-off, the more likely that there are changes in (unobserved) individual characteristics that may affect the outcome, such as health or household income. If these changes generally move in a specific direction over time, for example health deteriorates over time, estimation of the treatment effect will be biased.

On the other hand, the wider the window, the more precisely coefficients can be estimated, as more observations are used. Moreover, it is plausible to assume that in wider windows anticipation and adaptation effects are less pronounced since these are more likely to be observed in the vicinity of the cut-off.

Given the large sample sizes in all pre-sickness wage groups, the variance of the coefficient estimates is less of a concern than the potential bias introduced if the LR assumptions do not hold. Even in the smallest possible window there are some 5,000 to 25,000 observations in the data for each pre-sickness wage group. Therefore, the smallest possible window of 2 months around the cut-off, which is  $\{-1, 0\}$ , could be selected. However, for all analyses, we select the window of 24 months since smaller windows more often reflect anticipation and adaptation effects of the financial incentives (Section 4). In the appendix we present results for all windows from 2 to 48 months to show sensitivity.

In the LR design, potential outcomes cannot depend on the running variable inside the window  $W_0$  due to the exclusion restriction assumption. In the appendix we provide statistical evidence for this assumption for windows of 10 to 22 months for labor participation, and for all windows for benefit receipt from three alternative social support programs. The potential outcomes can, however, depend on the running variable outside the window  $W_0$ . This is because the exclusion restriction assumption does not need to hold outside this window. This, however, also means that the LR estimates outside the window do not need to reflect causal effects. Therefore, we take the long-term effect estimates as suggestive of causal effects.

## 5.4 Regression specification

We estimate the following regression model:

$$y_{ir} = \alpha_i + \beta T_{ir} + X_{ir}\gamma + \varepsilon_{ir}. \quad (1)$$

$i$  indexes individuals.  $r$  indexes the event time. Its values from  $-60$  to  $-1$  indicate the months before individuals face the financial incentives (first stage of the DI scheme),  $0$  is the cut-off month when first subjected to the financial incentives of the DI scheme (first month of the second stage of the DI scheme), and  $1$  to  $92$  indicate the remaining months of participation in the scheme.

$y$  is an outcome variable. It is labor participation or benefit receipt (social minimum supplement, social assistance, or unemployment insurance). The coefficient  $\beta$  on the treatment indicator  $T$  is the main parameter of interest. Assuming that the identifying assumptions hold, it captures the mean effect of the financial incentives around the cut-off.  $X$  is the vector of covariates related to the running variable. It includes dummies controlling for multiples of six months of DI receipt duration, and contemporary job vacancy rate. The dummies capture the potentially strong relationship between labor participation and DI duration. The job vacancy rate captures macroeconomic shocks.<sup>11</sup>  $\alpha_i$  is an individual-specific constant. It captures time-invariant labor participation or benefit claiming differences across individuals. It also captures differences in pre-sickness wage levels or time-invariant health conditions.  $\varepsilon_{ir}$  is an idiosyncratic (unobserved) shock, assumed to be uncorrelated with all explanatory variables.

Only observations for which the value of the running variable falls within the window,  $r \in W_0$ , are used in the estimations. Unless otherwise stated, we maintain the same regression specification across all analyses. The confidence intervals for coefficient estimates use standard errors that account for heteroskedasticity and clustering at the individual level in all regressions.

## 6 Effects of the financial incentives on labor participation and benefit receipt

### 6.1 Baseline effects

The top-left panel of Figure 3 presents the estimated effects of the financial incentives on labor participation considering a window of 24 months around the cut-off month when financial incentives take effect. The effects are sizable which suggest that financial incentives are effective in inducing work resumption in all pre-sickness wage groups.<sup>12</sup> The effects also exhibit a wage gradient. Higher pre-sickness wage earners more often resume work. This is due to the design of the financial incentives that are larger if pre-sickness wages are larger which imply larger income effects for higher pre-sickness wage earners who cannot manipulate incentive sizes. Those who earn at most the minimum wage increase labor participation by about 2 pp while those who earn at least 2.5 times of the minimum wage increase labor participation by about 5 pp after facing the financial incentives for one year.

The estimated effects might be attenuated by health status. Higher pre-sickness wage groups may be more able to invest in health care and recover from sickness compared to the lower pre-

---

<sup>11</sup>We cannot control for calendar time and age because they are perfectly collinear with the dummies for DI duration. We consider that the job vacancy rate captures the macroeconomic shocks that calendar time would capture. To investigate the possible effect of age on the treatment, we conduct heterogeneity analysis with respect to age.

<sup>12</sup>In the regressions, the dummies controlling for multiples of six months of wage-related benefit duration are almost always significant except in the smallest window and for the lowest pre-sickness wage group. The job vacancy rate is always insignificant.

sickness wage groups. In Section 6.3, we show that health does not exhibit a discontinuous change at the cut-off month when financial incentives take effect or at later months, suggesting that health does not affect the responses to the financial incentives.

The top-right and the bottom panel of Figure 3 present the estimated effects of the financial incentives on receipt of benefits from alternative social support programs. The results for social minimum supplement and social assistance show immediate responses that remain stable across all windows. Claims of social minimum supplement increase by large amounts of 4 pp among those who earn at most the minimum wage and by 9 pp among those who earn at least 2.5 times of the minimum wage, after facing the financial incentives for one year. This suggests that some individuals struggle to resume work after they exhaust their wage-related benefit, and substitute wage income with income from social support programs. This finding is in line with the studies that show that DI reforms that limit DI eligibility induce work resumption but also participation in alternative social support programs to compensate for the lost DI benefits (Karlström et al., 2008; Staubli, 2011; Borghans et al., 2014; Kantarci et al., 2023). The insignificant effect on social assistance for the lowest wage group is not surprising. As the bottom-left panel of Figure 2 showed, this group receives social assistance much more often than other groups both before and after facing the financial incentives. They are therefore much less likely to make additional claims of social assistance when they face the financial incentives.

The bottom-right panel of Figure 3 shows small and mostly insignificant effects for UI. This is expected since individuals exhaust their UI entitlement during their receipt of the wage-related benefit. One pre-sickness wage group shows a significant effect. As discussed in Section 4, it might be that they resume work during the wage-supplement or the follow-up benefit period but lose their job after a while and become entitled to UI.

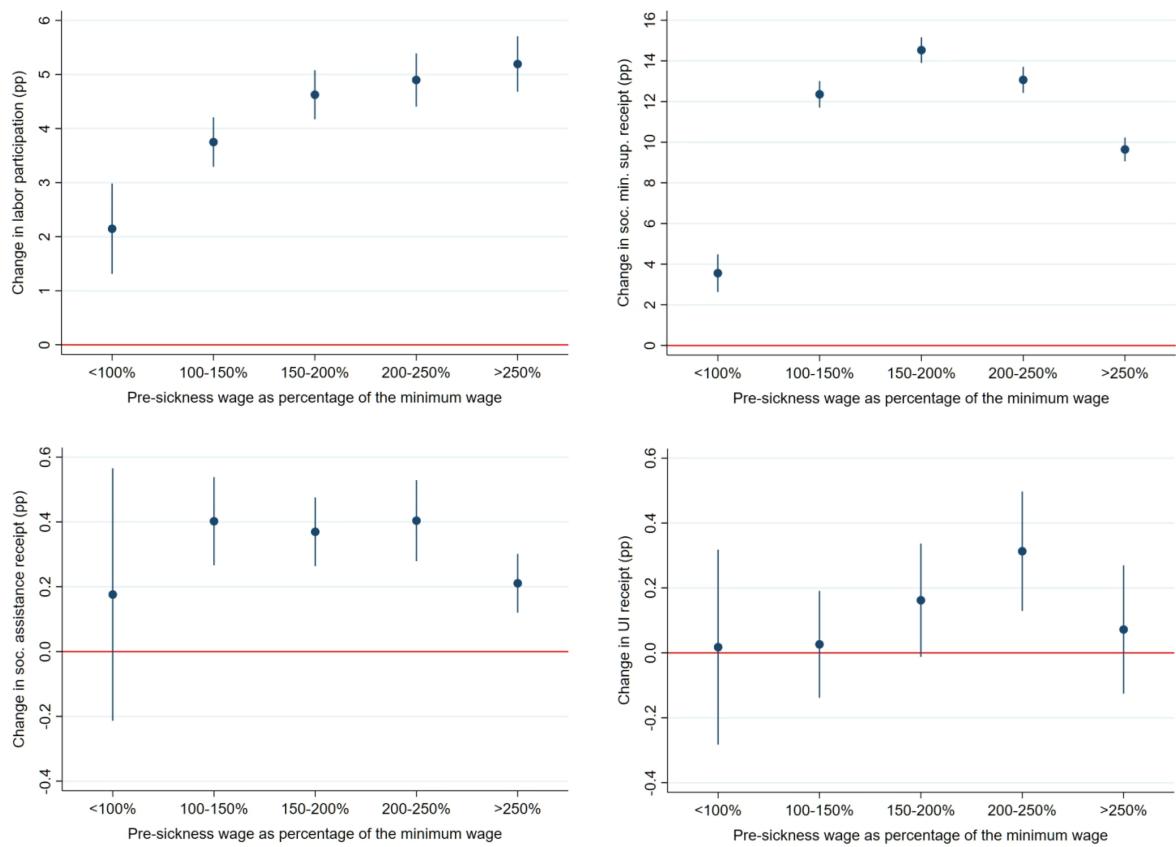


Figure 3: Estimated effects of the financial incentives on labor participation and receipt of three types of benefits and 95% confidence intervals around them.

## 6.2 Heterogeneous effects

We analyze the heterogeneous effects of the financial incentives with respect to the business cycle, duration of benefit receipt, and contract type (permanent contract, temporary contract, unemployment). It is, however, difficult to identify the heterogeneous effects because they can be correlated with other observable and unobservable characteristics. For example, for identifying the impact of benefit duration, individuals with shorter (wage-related) benefit durations will have shorter work histories (Section 2) and therefore will be younger on average. Their differential labor supply responses with respect to heterogeneous benefit durations can then be attenuated by an age effect. To address this, we consider a set of covariates as main variables driving compositional differences across heterogeneous groups of a factor. We weigh groups to have similar distributions of these covariates. We generate the group weights using entropy balancing (Hainmueller, 2012). Individuals are weighted to adjust inequalities in representation with respect to the first moment of the covariate distributions. As covariates we consider age, gender, contract type, and year of facing the financial incentives at the cut-off. We then estimate equation (1) within each heterogeneous group of a factor using the weights for that group.

### Business cycle

Impact of the financial incentives may vary over time for several reasons. Employment opportunities for disabled individuals are already limited, and these limitations can become more pronounced during economic downturns. This can in fact explain why Koning and van Sonsbeek (2017) found small effects of the financial incentives on labor participation, despite the incentives being large. Their study looked at the period 2006-2013 which for most individuals coincided with the financial crisis and higher unemployment. The impact of the financial incentives might also change over time due to factors endogenous to how beneficiaries manage financial incentives. Financial literacy of DI recipients may improve over time or DI awardees of later years may have access to better information of the financial incentives and therefore develop a better understanding of their implications for their income and respond more strongly.

Figure 4 presents the estimated effects of the financial incentives across cohorts of individuals with respect to the calendar year in which they exhausted their wage-related benefit and faced the financial incentives. We pool cohorts from three consecutive years. The cohorts who faced the incentives in 2006-2008, 2009-2011, 2012-2014, and 2018-2020 are re-weighted using entropy balancing so that cohorts have similar distributions of age, gender, and contract type. The impact of the financial incentives on labor participation is smallest for the 2009-2011 cohort and increases monotonically for cohorts who faced the financial incentives in subsequent years. Claims of the social minimum supplement and social assistance plausibly increase or decrease when labor participation decreases or increases, respectively. The business cycle can explain these patterns. Figure A4 in the appendix shows the business cycle indicator constructed by the Dutch Central Bank to identify turning points in the Dutch business cycle (Butler et al., 2019).<sup>13</sup> With a lag of about one year, the turning points of the indicator appear to predict well the differential labor supply responses across the cohorts.

To analyze the impact of the business cycle, we lag the business cycle indicator by one year and evaluate the impact of the financial incentives at two values of the indicator: when it is -1.5 and 1.5, which correspond well to the turning points of the indicator in Figure A4. In particular, we add an interaction term of the treatment dummy and the lagged business cycle indicator in equation (1), and evaluate the marginal effect of the treatment dummy and its interaction

<sup>13</sup>The indicator is composed of 86 potential sub-series that are closely related to the development of real GDP growth. Figure A4 shows that the indicator is capable of identifying tipping points in real GDP growth.

with the lagged business cycle indicator at these two values of the indicator. Figure 5 presents the estimation results. The impact of the financial incentives on labor participation is larger during economic upturns than during downturns for all pre-sickness wage groups. The effect is less pronounced for the lowest wage group suggesting that labor market opportunities improve only to a limited extent for this group during economic upturn. The impact of the financial incentives on receipt of alternative benefits is in line with the impact on labor participation. Facing the financial incentives during economic downturn, some partially disabled individuals cannot resume work and instead turn to benefits from social support programs.

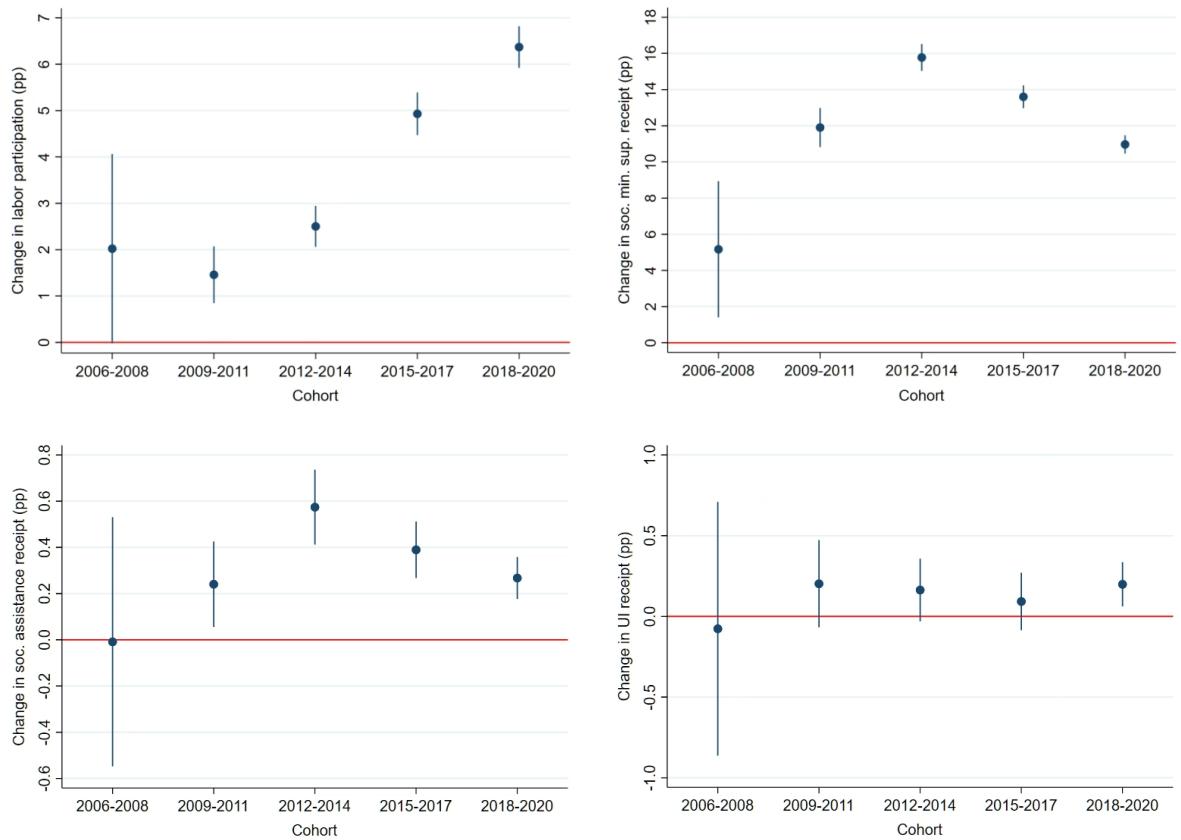


Figure 4: Estimated effects of the financial incentives on labor participation and receipt of three types of benefits and 95% confidence intervals around them for cohorts of individuals with respect to the calendar year they exhausted their wage-related benefit.

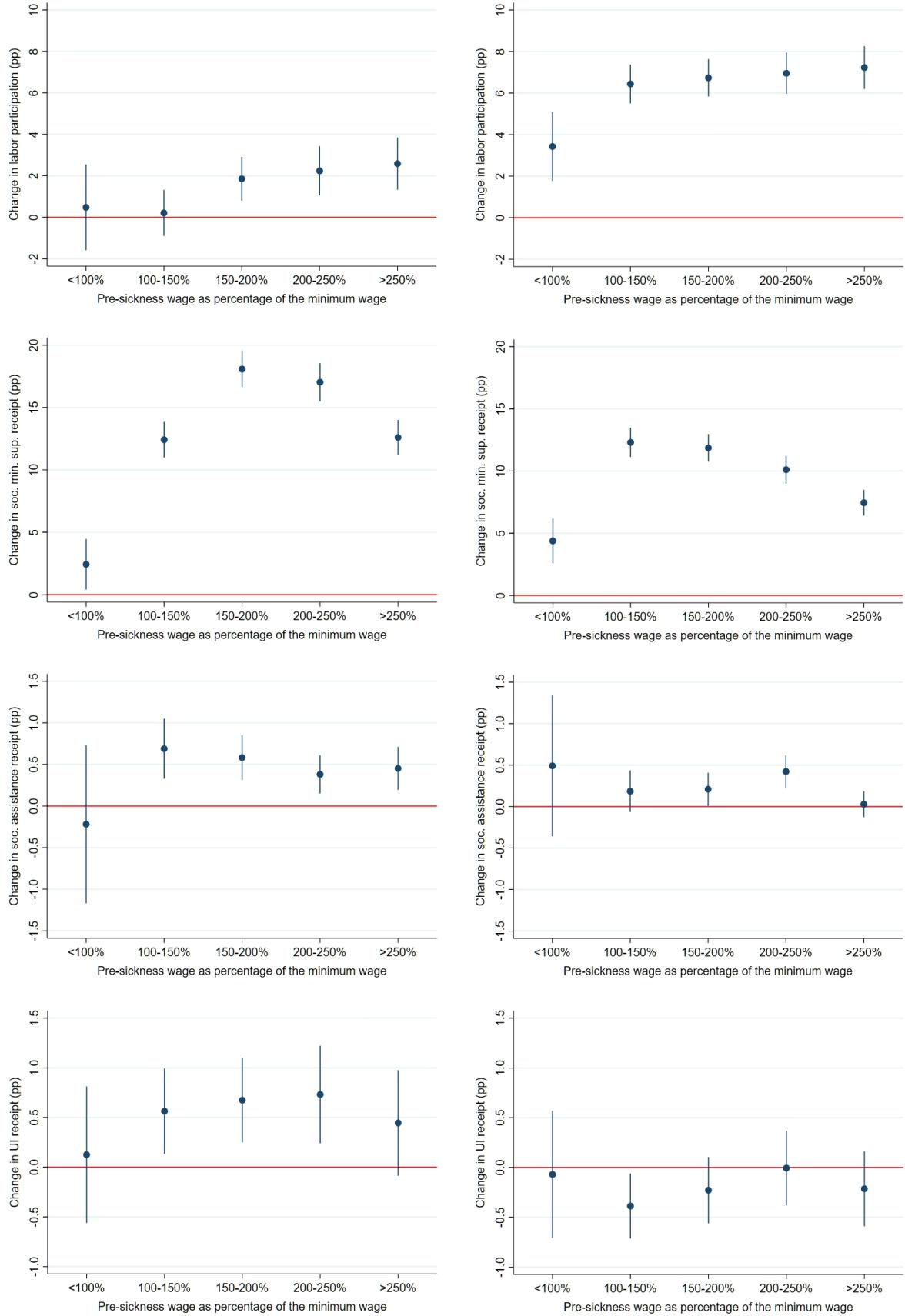


Figure 5: Estimated effects of the financial incentives on labor participation and receipt of three types of benefits and 95% confidence intervals around them for individuals who faced the financial incentives when the business cycle index was  $-1.5$  (left panel) and when it was  $1.5$  (right panel).

## Duration of benefit receipt

A large number of studies analyze which policy measures are effective to limit inflow into DI or stimulate outflow from DI (Section 1). Some of these studies compare policy measures and argue that those that tighten the criteria to enter DI schemes are more effective than those that stimulate reintegration in DI schemes (Van Sonsbeek and Gradus, 2013; Haller et al., 2024; Kantarci et al., 2023). These studies, however, do not account for the role of sickness duration. Labor market attachment and health status can be very different when people apply for DI the first time and after they spend some time claiming DI. A particular policy measure may therefore be effective at the onset of sickness but lose its effectiveness at later stages of sickness. For example, longer disability periods may lead to more human capital loss or a stronger scarring effect (Arulampalam, 2001; Arulampalam et al., 2001), reducing the chances of responding to financial incentives to resume work over time. This implies that as different policy measures typically apply at different stages of sickness, sickness duration can affect the assessment of which measure is more effective than another to stimulate work resumption.

In the WGA, the number of months individuals spend claiming the first-stage wage-related benefit, and hence when they face the financial incentives at the cut-off month, is individual specific (Section 2). Figure A2 in the appendix shows the substantial individual variation in the duration of the wage-related benefit. Moreover, this variation can be treated as random since individuals cannot manipulate the duration of the wage-related benefit, which depends on their work history. We exploit this random variation in timing of the financial incentives and analyze how the effectiveness of the financial incentives depends on benefit duration.

Figure 6 presents the estimated effects of the financial incentives for two groups of individuals who have spent at most and at least 2 years claiming the wage-related benefit. The group of individuals with at most 2 years of benefit duration is re-weighted using entropy balancing so that the two duration groups have similar distributions of age, gender, contract type, and year of facing the financial incentives at the cut-off. The effects of the financial incentives on labor participation and benefit receipt in both groups are smaller than the effects found in baseline analysis. This is due to entropy balancing where the group with shorter benefit duration is matched in observable covariates to the group with longer benefit duration. The latter group is relatively older and more often has a permanent contract and exhibit smaller responses. We find that the responses of the two groups are similar. This suggests that effectiveness of the financial incentives to induce work resumption does not depend on the time spent claiming DI.

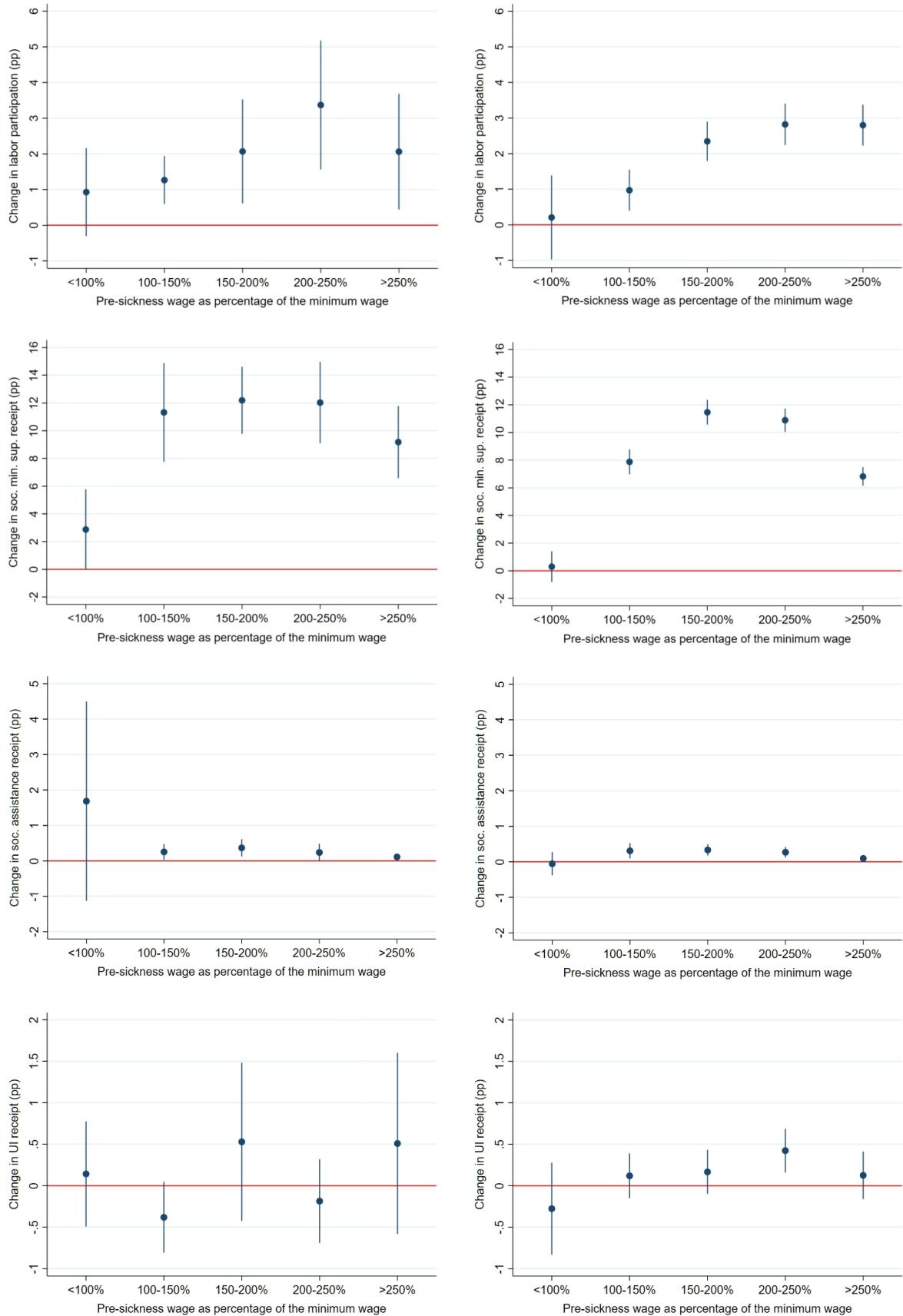


Figure 6: Estimated effects of the financial incentives on labor participation and receipt of three types of benefits and 95% confidence intervals around them for individuals who spend at most (left panel) and at least (right panel) 2 years claiming the wage-related benefit.  
25

## Contract type

Kantarcı et al. (2023) and Bernasconi et al. (2024) analyzed the effects of stricter eligibility criteria and reduced benefit generosity due to the Dutch DI reform in 2006. They find that the reform led to an increase of labor participation of the individuals who fell sick only if these individuals had a permanent job, whereas spouses responded to the DI reform if individuals reporting sick had a weak labor market position in terms of working in temporary contract jobs or being unemployed. The impact of the financial incentives could also exhibit heterogeneity across employees with a permanent contract and those with a temporary contract and the unemployed. As explained in Section 2, the financial incentives to resume work for partially disabled individuals are strong, but they are offered only after a period of up to about 4 to 5 years after the first day of sickness (Section 2), complicating a successful return to the labor market. During the sickness period, individuals employed through temporary work agencies and the unemployed have no former employer to return to and therefore carry a higher risk of unemployment than those with permanent contracts when they face the financial incentives. In fact, in our data, 6 months prior to facing the financial incentives, labor force non-participation rate is 41% for permanent contract workers whereas it is 74% and 79% for temporary contract workers and the unemployed, respectively. If disabled individuals do not have a job when they face the financial incentives, they may struggle to resume work if job search costs are high, negotiating a suitable work schedule with an employer is difficult, or if the offered wage in the new job is lower than the reservation wage (Koning and Lindeboom, 2015; Zaresani, 2020). For these reasons, financial incentives may be less effective among temporary contract workers and the unemployed. On the other hand, the much lower labor participation rates among the temporary contract workers and the unemployed prior to facing the financial incentives may imply substantial unused work capacity. These groups may therefore be more responsive to the financial incentives.

Figure 7 presents the estimated effects of the financial incentives by contract type. Groups of individuals who have a temporary contract and those who are unemployed are re-weighted using entropy balancing so that the three labor market groups have similar distributions of age, gender, and year of facing the financial incentives at the cut-off. The labor participation responses of the temporary contract workers and the unemployed are much stronger than those of the permanent contract workers, especially in higher pre-sickness wage groups where the financial incentives are stronger. As explained above, the potential reason is that the labor participation rate is lower for these groups than for permanent contract workers when they face the financial incentives at the cut-off date meaning that there is more room to increase labor participation in these groups. However, these groups also more often claim benefits from alternative social support programs when they face the financial incentives, possibly because some are not able to find a job despite the financial incentives and have to rely on alternative benefits when they lose DI benefits. The results for the unemployed seem to confirm this. As job opportunities are likely to be limited for this group more than for temporary contract workers, possibly due to human capital loss and a scarring effect, they become more reliant on alternative social support programs when they lose DI benefits.

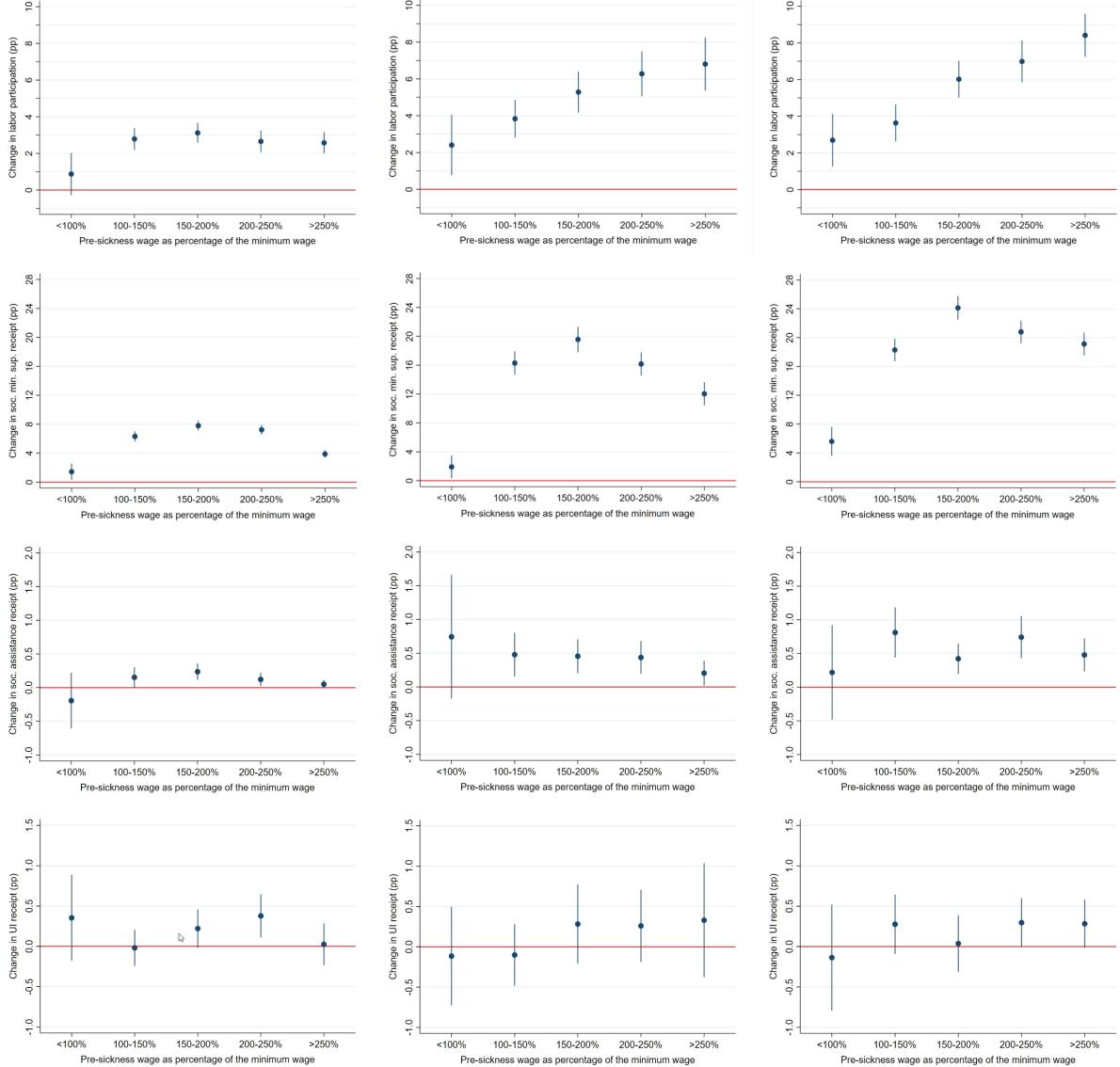


Figure 7: Estimated effects of the financial incentives on labor participation and receipt of three types of benefits and 95% confidence intervals around them for individuals who have a permanent contract (left panel), temporary contract (center panel), and unemployed (right panel).

### 6.3 Checking the identifying assumptions

In this section we conduct falsification tests to provide evidence in support of the LR approach to RD, but also estimate models taking alternative identification strategies. We relegate the details and extended results to the online appendix.

The LR approach assumes that the potential outcomes do not depend on the value of the running variable inside a window except via the treatment assignment indicator. We check if the treatment effect is insensitive to window choice which implies that potential outcomes do not depend on the value of the running variable in that window. We find statistical evidence for this assumption. For labor participation, the difference between the treatment effect estimates from windows of 12 to 22 months and the treatment effect estimate from the benchmark window of 10 months are statistically indifferent from each other (at 10%) in all pre-sickness wage groups. For benefit receipt from alternative social support programs, treatment effect estimates from all windows are statistically indifferent from each other in all pre-sickness wage groups. Extended results are presented in Section

In a RD design, there should be no treatment effect at placebo cut-offs. We consider windows of 24 months around placebo cut-offs that precede and succeed the true cut-off by 18 months. For labor participation, we find statistically significant effects at the placebo cut-off month of -18. The magnitudes of these effects are substantially smaller than the baseline effect at the true cut-off. We attribute this placebo effect to work resumption regardless of the financial incentives and to anticipation due to the financial incentives. With rare exceptions, we find no placebo effect for any of the three types of benefit receipt.

In a RD design, relevant covariates should not discontinuously change at the cut-off. We consider health as the main relevant covariate. We measure health by the number of medicines prescribed in a year under the mandatory health care insurance act. We find that the number of medicines prescribed does not exhibit a discontinuity at the cut-off in any pre-sickness wage group or window except in rare cases.

As an alternative identification strategy, we consider the continuity-based approach to RD. We find that, in the smallest window of 2 months, the results based on the LR approach and the continuity-based approach lead to similar estimates of the treatment effect for all outcomes. The estimates are also similar when we consider donut-hole regressions. We repeated the falsification tests above for the continuity-based approach to RD. Unlike in the LR approach, we find no significant effect at placebo cut-offs. As in the LR approach, the number of medicines prescribed does not exhibit a discontinuity at the cut-off in any pre-sickness wage group.

Koning and van Sonsbeek (2017) analyzed the impact of the financial incentives of the WGA using data from 2006 to 2013. We estimate their baseline model, and find that their and our identification strategies lead to similar treatment effect estimates.

### 6.4 Full treatment effect

Identification of the true effect of the financial incentives requires that individuals do not change work and benefit claiming decisions in anticipation of the incentives before they face the incentives, or do not take time to respond to the incentives due to adaptation after they face the incentives the first time at the cut-off (Section 5.2). In Figure 2, notable increases in labor participation in the vicinity of the cut-off month of the financial incentives suggested anticipation and adaptation effects. Assuming that anticipation and adaptation are part of the true responses, we investigate to which extent they lead to an underestimation of the true effect of the financial incentives in the baseline regression. Following Shigeoka (2014), we consider a donut hole regression to net out the effects of anticipation and adaptation and estimate the true effect of the incentives. Our donut hole is the window of 8 months around the cut-off month.

This window is chosen based on the observation that, in Figure 2, work resumption shows a gradual increase in windows of 2 to 8 months around the cut-off for all wage groups. Figure 8 presents the estimation results. For labor participation, the estimates are about twice as large as the baseline estimates in Figure 3. This suggests that the baseline specification underestimates the effect of financial incentives in work resumption due to anticipation and adaptation effects. Therefore, the baseline estimates can be viewed as lower bounds to the full treatment effects of the financial incentives. For receipt of three types of benefits, we find small differences between the donut-hole and baseline estimates.

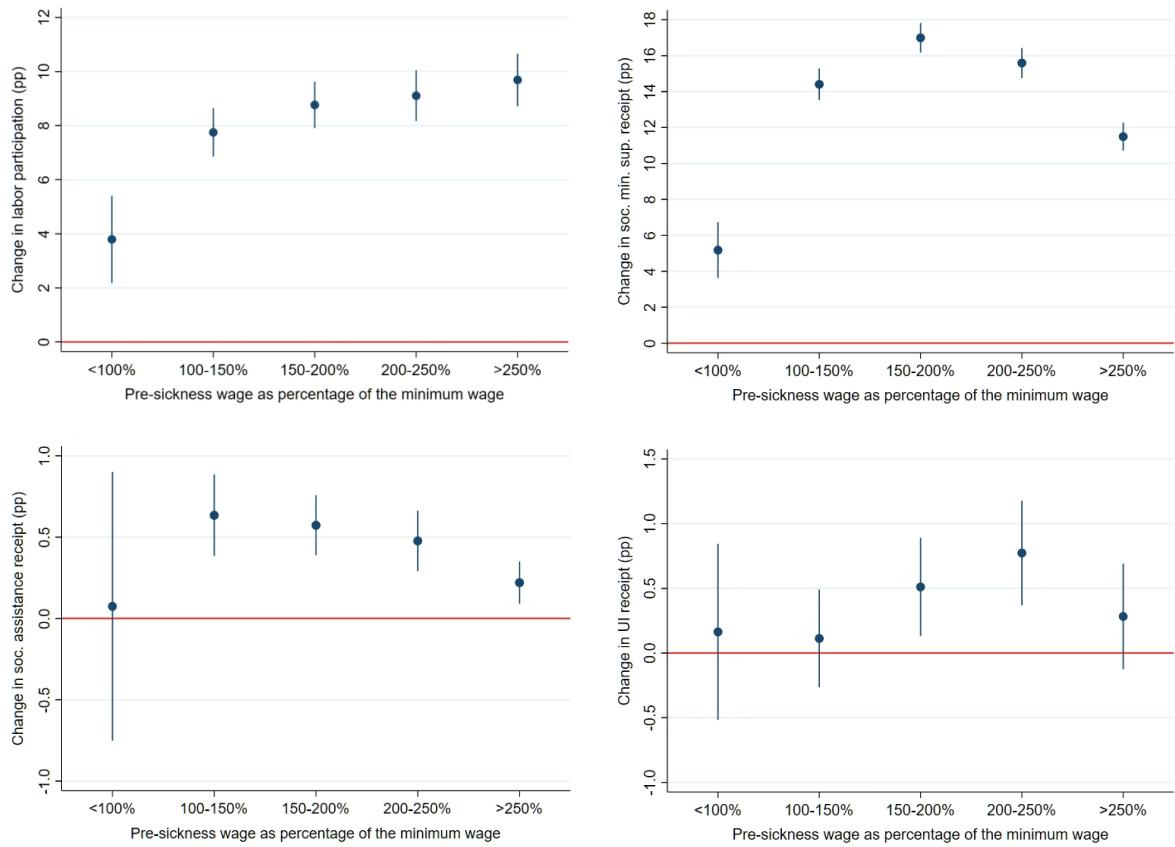


Figure 8: Estimated effects of the financial incentives on labor participation and receipt of three types of benefits and 95% confidence intervals around them based on donut hole regressions.

## 6.5 Elasticity of labor force non-participation

The baseline results suggest that partially disabled individuals have substantial remaining work capacity and financial incentives are effective to induce them to utilize it. As discussed in Section 2, the disability benefit is reduced and the reduction is discontinuous at the cut-off month when the wage-related benefit expires. Moreover, the size of the reduction exogenously depends on the earnings of individuals before they fall sick. We exploit these design features to estimate the elasticity of labor force non-participation to the benefit reduction. We calculate the elasticity as the ratio of the observed change in labor force non-participation from the first-to the second-stage of the DI scheme to the “potential” change in income over the same period. The latter change, that is, the denominator, means that we consider individuals as we observe them at the end of the first stage of the scheme, and calculate their potential financial incentive to work in the second stage. We measure the changes in percentages.

We calculate the observed change in labor force non-participation as the difference between the fractions of those in non-employment in the first- and second-stages of the DI scheme. In the first-stage, we measure non-participation at (event) month  $-5$ , that is, 5 months preceding the cut-off month when the financial incentives become effective. In the second-stage, we measure non-participation over the first 24 months of the second-stage.

We calculate the potential change in income as follows. Income in the first-stage is the income from wage (if individuals work) and the wage-related benefit as we observe them. The wage-related benefit in the first stage depends on the pre-sickness wage, disability grade, and the remaining earning capacity utilization rate. Pre-sickness wage is given. We calculate the remaining earning capacity utilization rate as the ratio of earnings to remaining earning capacity. The latter is given by the difference between the pre-sickness earnings and pre-sickness earnings multiplied by the disability grade. In the first-stage, we calculate the income at event month  $-5$ . In the second stage, we assume that individuals do not work and their income remains constant and is equal to the follow-up benefit. We also assume that individuals do not expect their disability grade to change and hence it remains constant in the second stage of the scheme.

We find that the elasticity of labor force non-participation is 0.2. This compares to the estimates of [Kostøl and Mogstad \(2014\)](#) that range between 0.1 and 0.3 over the period 2005-2007 in Norway. Our estimate is larger than those of [Ruh and Staubli \(2019\)](#) and [Zaresani \(2020\)](#), which are 0.1 and 0.11 in Austria and Canada, respectively. Table 2 presents the elasticity by pre-sickness wage group. Consistent with our baseline estimates of the impact on labor participation, higher pre-sickness wage groups exhibit larger elasticities. This is expected as the potential incentive to resume work is larger for these groups (see the third column in Table 2). Note that the elasticities are larger for higher pre-sickness wage groups despite that labor force non-participation rates prior to facing the financial incentives are smaller for these groups (fourth column in Table 2) which may imply less unused work capacity. This shows that financial incentives are a strong determinant of work resumption. This highlights the significant role of financial incentives in determining work resumption.

Table 2: Elasticity of labor force non-participation by pre-sickness wage group

Pre-sickness wage as a fraction of the minimum wage	Elasticity of labor non-participation	Potential daily incentive (€)	Labor non-participation rate at t = -5 (%)
<100%	0.05	27.7	64.5
100-150%	0.18	46.1	60.3
150-200%	0.20	72.2	55.8
200-250%	0.22	99.2	54.8
>250%	0.24	138.4	50.0

## 7 Conclusion

Targeting benefit recipients in a DI scheme is important as a policy intervention because effects which operate on the existing stock of beneficiaries have the potential to make a much greater immediate impact than effects which operate on the comparatively small inflow into the DI scheme (Borghans et al., 2014). A DI scheme dedicated to partially disabled individuals, in particular, can accommodate labor market policies focusing on activating remaining work capacity as in the Netherlands (De Jong and de Vos, 2005). Partially disabled individuals, however, can engage in moral hazard and claim disability benefits for the part of their remaining work capacity they do not utilize. While the Employee Insurance Agency examines all sick-listed workers of their remaining work capacity at the end of the SI period, changes in work capacity remain largely unexamined during the succeeding DI period because the Employee Insurance Agency can conduct only a limited number of reexaminations due to the shortage of medical examiners but also the obvious cost of reexamining large numbers of beneficiaries. Moreover, it is difficult to observe if partially disabled individuals actively search for work and fail to receive job offers to deserve the DI benefits they receive. Financial incentives can be an effective measure to induce work resumption at least because they do not necessarily involve reexamination protocol.

We analyzed the effects of financial incentives built into the Dutch DI scheme for partially disabled individuals. We find that, on average, financial incentives are effective in inducing work resumption. For the average wage earner (before falling sick) in our population, the financial incentives increase labor participation by 4.5 pp. Netting out the anticipation and adaption effects of the incentives, this estimate increases to about 9 pp. This suggests that partially disabled individuals indeed have considerable remaining capacity to work. In fact, compared to permanent contact workers, the effects are larger among temporary contract workers and the unemployed as labor market groups deemed to have limited job opportunities and be vulnerable. This suggests that moral hazard is likely and financial incentives are effective to insure against it.

We find that financial incentives are much less effective to induce work resumption among more than a quarter of DI recipients who are low-wage earners. By design, their financial incentive to resume work is comparatively small, and, on top of their health impairment, lack of skills may put additional constraints on their job opportunities. In fact, we find that, instead of resuming work, they much more often rely on benefits from alternative social support programs. Although high-wage earners have more to lose from the stronger financial incentives, due to the larger income effect of the financial incentives for them, they are also better able to compensate a disability related income loss through alternative income sources. This shows that the financial incentives may unintentionally increase inequality.

We find that the business cycle matters for the effectiveness of the financial incentives. This underlines the weak position of the disabled individuals on the labor market. In a recession and a loose labor market, their chances to find a job decrease significantly. This finding also suggests that evaluations of reintegration programs should account for the macroeconomic conditions.

## References

- Arulampalam, W., 2001. Is unemployment really scarring? Effects of unemployment experiences on wages. *The Economic Journal* 111 (475), F585–606.
- Arulampalam, W., Gregg, P., Gregory, M., 2001. Introduction: unemployment scarring. *The Economic Journal* 111 (475), F577–584.
- Bernasconi, M., Kantarcı, T., van Soest, A., van Sonsbeek, J.-M., 2024. The added worker

- effect: Evidence from a disability insurance reform. *Review of Economics of the Household* E-pub ahead of print.
- Borghans, L., Gielen, A. C., Luttmer, E. F. P., 2014. Social support substitution and the earnings rebound: evidence from a regression discontinuity in disability insurance reform. *American Economic Journal: Economic Policy* 6 (4), 34–70.
- Bound, J., Burkhauser, R. V., 1999. Chapter 51 economic analysis of transfer programs targeted on people with disabilities. Vol. 3 of *Handbook of Labor Economics*. Elsevier, pp. 3417–3528.
- Burkhauser, R., Marc, D., McVicar, D., Wilkins, R., 2014. Disability benefit growth and disability reform in the US: lessons from other OECD nations. *IZA Journal of Labor Policy* 3 (4), 1–30.
- Butler, B., Volkerink, M., Lung, C. L., 2019. Nieuwe conjunctuur indicator voorspelt afvlakking groei in 2019. *Economisch Statistische Berichten* 104 (4773).
- Campolieti, M., 2004. Disability insurance benefits and labor supply: some additional evidence. *Journal of Labor Economics* 22 (4), 863–889.
- Campolieti, M., Riddell, C., 2012. Disability policy and the labor market: evidence from a natural experiment in Canada, 1998–2006. *Journal of Public Economics* 96 (3-4), 306–316.
- Cattaneo, M. D., Frandsen, B. R., Titiunik, R., 2015. Randomization inference in the regression discontinuity design: An application to party advantages in the U.S. senate. *Journal of Causal Inference* 3 (1), 1–24.
- Cattaneo, M. D., Idrobo, N., Titiunik, R., 2020. *A Practical Introduction to Regression Discontinuity Designs: Foundations. Elements in Quantitative and Computational Methods for the Social Sciences*. Cambridge University Press.
- Cattaneo, M. D., Titiunik, R., 2021. Regression discontinuity designs. Mimeo.
- De Jong, P. R., de Vos, E. L., 2005. Lessons from the Dutch experience. *Revue française des affaires sociales* 1 (2), 183–205.
- Favre, G., Haller, A., Staubli, S., 2021. Induced entry in disability insurance: evidence from Canada. National Bureau of Economic Research, Center paper NB20-08.
- Gruber, J., 2000. Disability insurance benefits and labor supply. *Journal of Political Economy* 108 (6), 1162–1183.
- Hainmueller, J., 2012. Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20 (1), 25–46.
- Haller, A., Staubli, S., Zweimüller, J., 2024. Designing disability insurance reforms: tightening eligibility rules or reducing benefits? *Econometrica* 92 (1), 79–110.
- Kantarcı, T., van Sonsbeek, J.-M., Zhang, Y., 2023. The heterogenous impact of stricter criteria for disability insurance. *Health Economics* 31 (9), 1898–1920.
- Karlström, A., Palme, M., Svensson, I., 2008. The employment effect of stricter rules for eligibility for di: Evidence from a natural experiment in sweden. *Journal of Public Economics* 92 (10-11), 2071–2082.
- Koning, P., Lindeboom, M., 2015. The rise and fall of disability insurance enrollment in the Netherlands. *Journal of Economic Perspectives* 29 (2), 151–172.
- Koning, P., van Sonsbeek, J.-M., 2017. Making disability work? The effects of financial incentives on partially disabled workers. *Labour Economics* 47, 202–215.
- Kostøl, A. R., Mogstad, M., 2014. How financial incentives induce disability insurance recipients to return to work. *American Economic Review* 104 (2), 624–655.
- Malani, A., Reif, J., 2015. Interpreting pre-trends as anticipation: Impact on estimated treatment effects from tort reform. *Journal of Public Economics* 124, 1–17.
- Mattei, A., Mealli, F., 2017. Regression discontinuity designs as local randomized experiments. *Observational Studies* 3 (2), 156–173.

- Mullen, K. J., Staubli, S., 2016. Disability benefit generosity and labor force withdrawal. *Journal of Public Economics* 143, 49–63.
- Ruh, P., Staubli, S., 2019. Financial incentives and earnings of disability insurance recipients: evidence from a notch design. *American Economic Journal: Economic Policy* 11 (2), 269–300.
- Sales, A. C., Hansen, B. B., 2020. Limitless regression discontinuity. *Journal of Educational and Behavioral Statistics* 45 (2), 143174.
- Shigeoka, H., 2014. The effect of patient cost sharing on utilization, health, and risk protection. *American Economic Review* 104 (7), 2152–2184.
- Staubli, S., 2011. The impact of stricter criteria for disability insurance on labor force participation. *Journal of Public Economics* 95 (9-10), 1223–1235.
- Vall Castelló, J., 2017. What happens to the employment of disabled individuals when all financial disincentives to work are abolished? *Health Economics* 26 (S2), 158–174.
- Van Sonsbeek, J.-M., Gradus, R. H. J. M., 2013. Estimating the effects of recent disability reforms in the Netherlands. *Oxford Economic Papers* 65 (4), 832–855.
- Weathers, R. R., Hemmeter, J., 2011. The impact of changing financial work incentives on the earnings of social security disability insurance (ssdi) beneficiaries. *Journal of Policy Analysis and Management* 30 (4), 708–728.
- Zaresani, A., 2020. Adjustment cost and incentives to work: Evidence from a disability insurance program. *Journal of Public Economics* 188 (104223).

## Online appendix

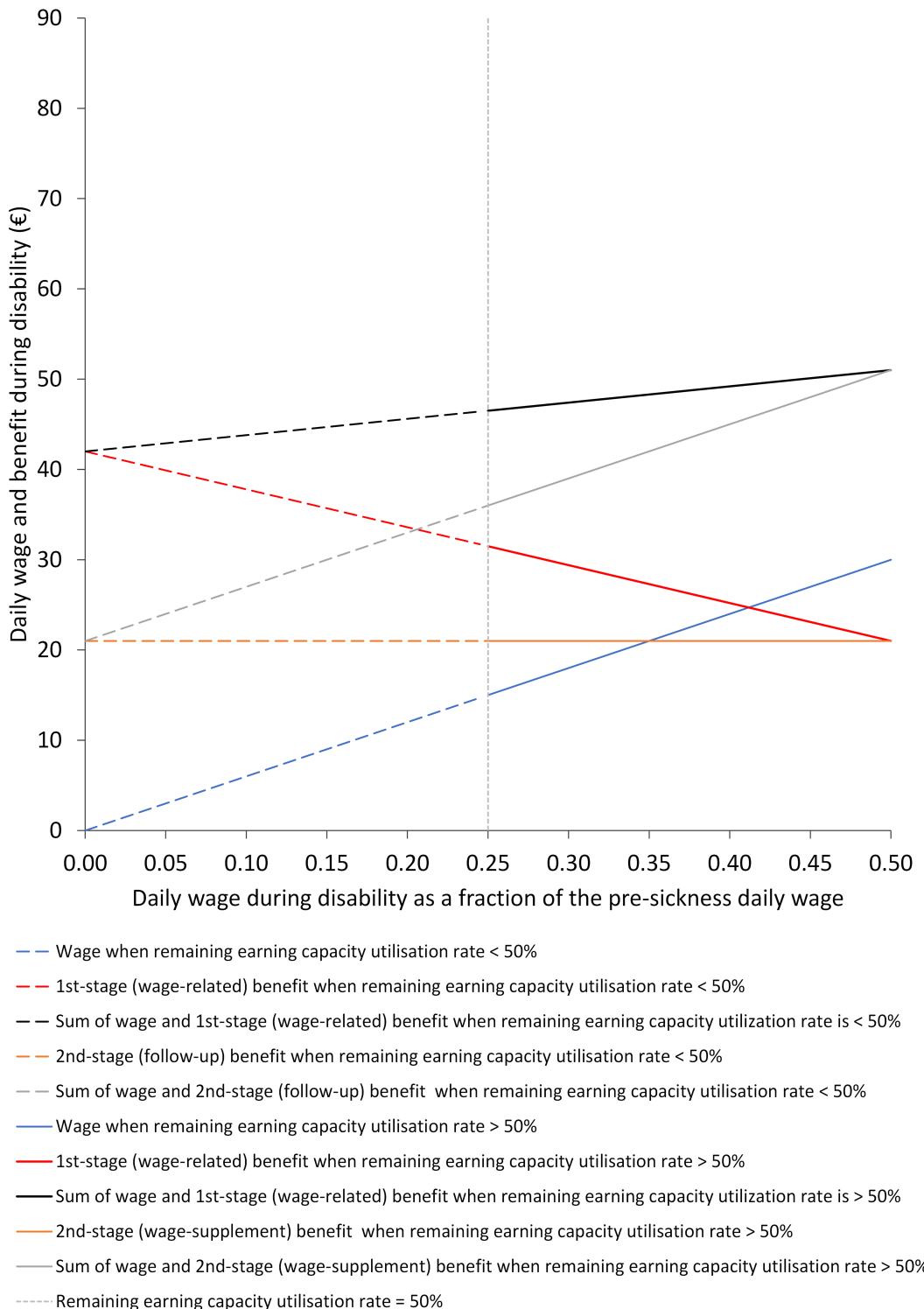


Figure A1: Financial incentives to work in the second stage of the DI scheme for an individual who earns €60 per day before falling sick and has a disability grade of 50%. Dashed lines apply when remaining work capacity utilization rate < 50%. Solid lines apply when remaining work capacity utilization rate  $\geq$  50%. The financial incentive is the difference between the dashed black line (sum of wage and the wage-related benefit) and dashed gray line (sum of wage and the follow-up benefit) if remaining work capacity utilization rate < 50% in the second stage of the DI scheme (left side of vertical reference line). The financial incentive is the difference between the solid black line (sum of wage and the wage-related benefit) and the solid gray line (sum of wage and the wage-supplement benefit) if remaining work capacity utilization rate  $\geq$  50% in the second stage of the DI scheme (right side of vertical reference line).

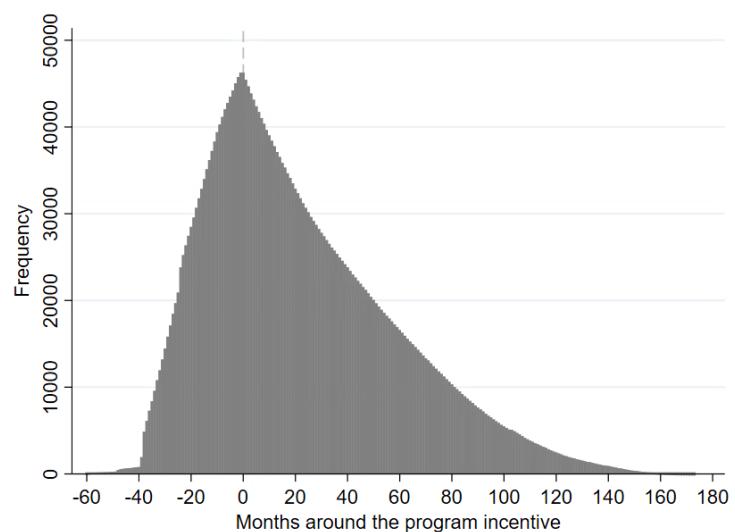


Figure A2: Number of months spent in the DI scheme before (first stage of the DI scheme) and after (second stage of the DI scheme) the month financial incentives take effect.

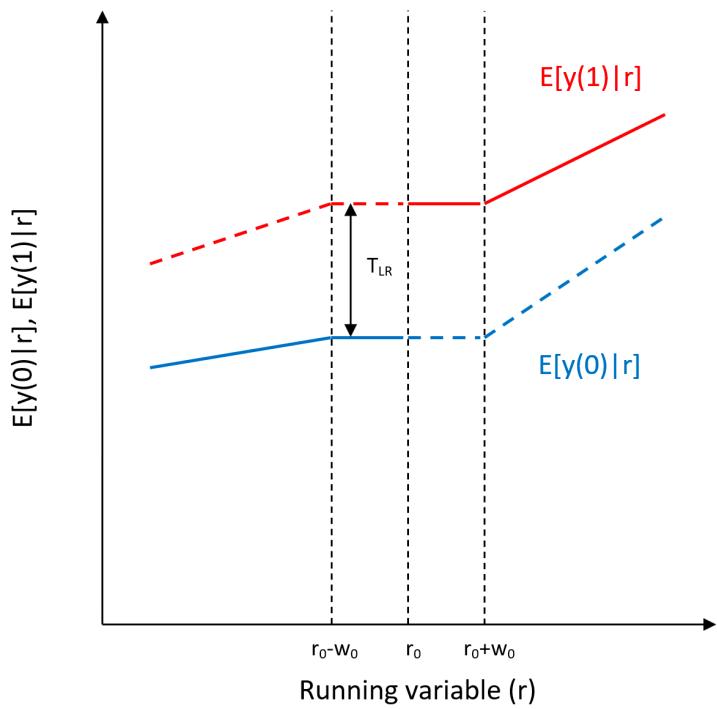


Figure A3: Illustration of the local randomization approach to RD based on [Cattaneo and Titiunik \(2021\)](#).

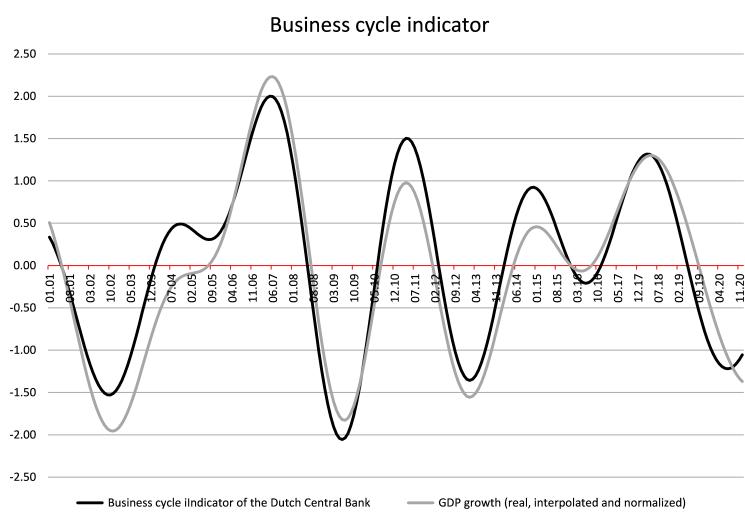


Figure A4: Business cycle indicator of the Dutch Central Bank.

## Checking the identifying assumptions

### Window selection

The exclusion restriction assumption of the LR method (Section 7) requires that the potential outcomes do not depend on the value of the running variable inside a window, except via the treatment assignment indicator (Figure A3).

Figure A5 shows that the baseline treatment effect estimate, as the estimate of the difference in the expected values of the potential outcomes before and after the cut-off, shows an increase in labor participation in the vicinity of the cut-off month when financial incentives take effect (windows of 2 to 8 months). As discussed in Sections 4 and 5.3, this is likely due to anticipation and adaptation because individuals are not able to respond to the financial incentives immediately at the cut-off month the incentives take effect. Besides the increase in the vicinity of the cut-off, however, the treatment effect estimate remains stable across most of the wider windows in all wage groups. Figure A6 shows that the baseline treatment effect estimates for benefit receipt from three social support programs are largely stable across all windows.

In statistical terms, for labor participation, the two-sample z-tests of the difference between the treatment effect estimates from windows of 2 to 6 months and the treatment effect estimate from a benchmark window of 10 months are statistically different from each other (at 5%) in all pre-sickness wage groups. This is due to anticipation and adaptation effects in the vicinity of the cut-off month when financial incentives take effect. However, the difference between the treatment effect estimates from windows of 12 to 22 months and the treatment effect estimate from the benchmark window of 10 months are not significant at 10% in all wage groups. These results suggest that the impact of the financial incentives takes full effect and stabilizes after an initial period of anticipation and adaptation responses that last about eight months. This provides evidence for the exclusion restriction assumption for windows of 10 to 22 months. For benefit receipt from three social support programs, the difference between the treatment effect estimate from the benchmark window of 10 and the treatment effect estimates from all other windows are statistically indifferent from each other except for two pre-sickness wage groups when we consider a window of 2 months.

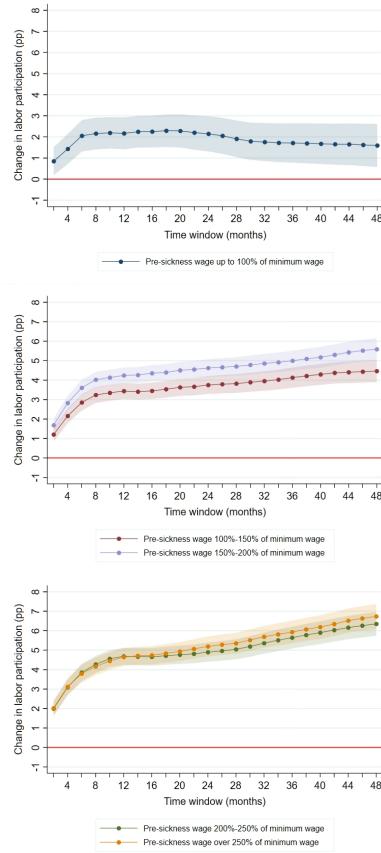


Figure A5: Estimated effects of the financial incentives on labor participation and 95% confidence intervals around them for different windows.

## Placebo cut-offs

Comparing individuals before and after they face the financial incentives, we showed significant effects of the financial incentives in the vicinity of the (true) cut-off when financial incentives become effective. To assess that these effects are caused by the treatment and not by other factors, we compare the mean outcome around placebo cut-offs, before and after the true cut-off, keeping otherwise the baseline regression specification the same. In particular, we consider placebo cut-offs at months -18 and 18 before and after the true cut-off, respectively, considering a window of 24 months around each placebo cut-off. At the placebo cut-off before the true cut-off, this leads to a comparison that is between individuals who have not yet been treated. At the placebo cut-off after the true cut-off, the comparison is between treated individuals. Significant differences around the placebo cut-offs may cast doubt on that the differences around the true cut-off are due to the treatment.

Figure A22 presents estimation results. For labor participation, we find significant effects at the placebo cut-off at month -18. However, the magnitudes of the placebo effects (top-left panel) are substantially smaller than the baseline effects at the true cut-off (top-center panel). We find no significant effects at the placebo cut-off at month 18. Apparently employment prospects or labor market attachment are stronger at the early stages of disability but they become weaker over time. This decreasing work resumption trend, however, is interrupted by a notable increase in work resumption at the true cut-off due to the financial incentives. In Figure A23 we show the transitions between working and not working in adjacent months around the (true) cut-off which confirm these trends. The transitions suggest that the significant placebo effects are part of work resumption during the early stages of DI participation regardless of the financial incentives. They could, however, also reflect early anticipation effects in response to the financial incentives.

We find almost no significant placebo effect for benefit receipt from alternative social support programs. These results do not cast serious doubt on our main finding that financial incentives change behaviour at the true cut-off.

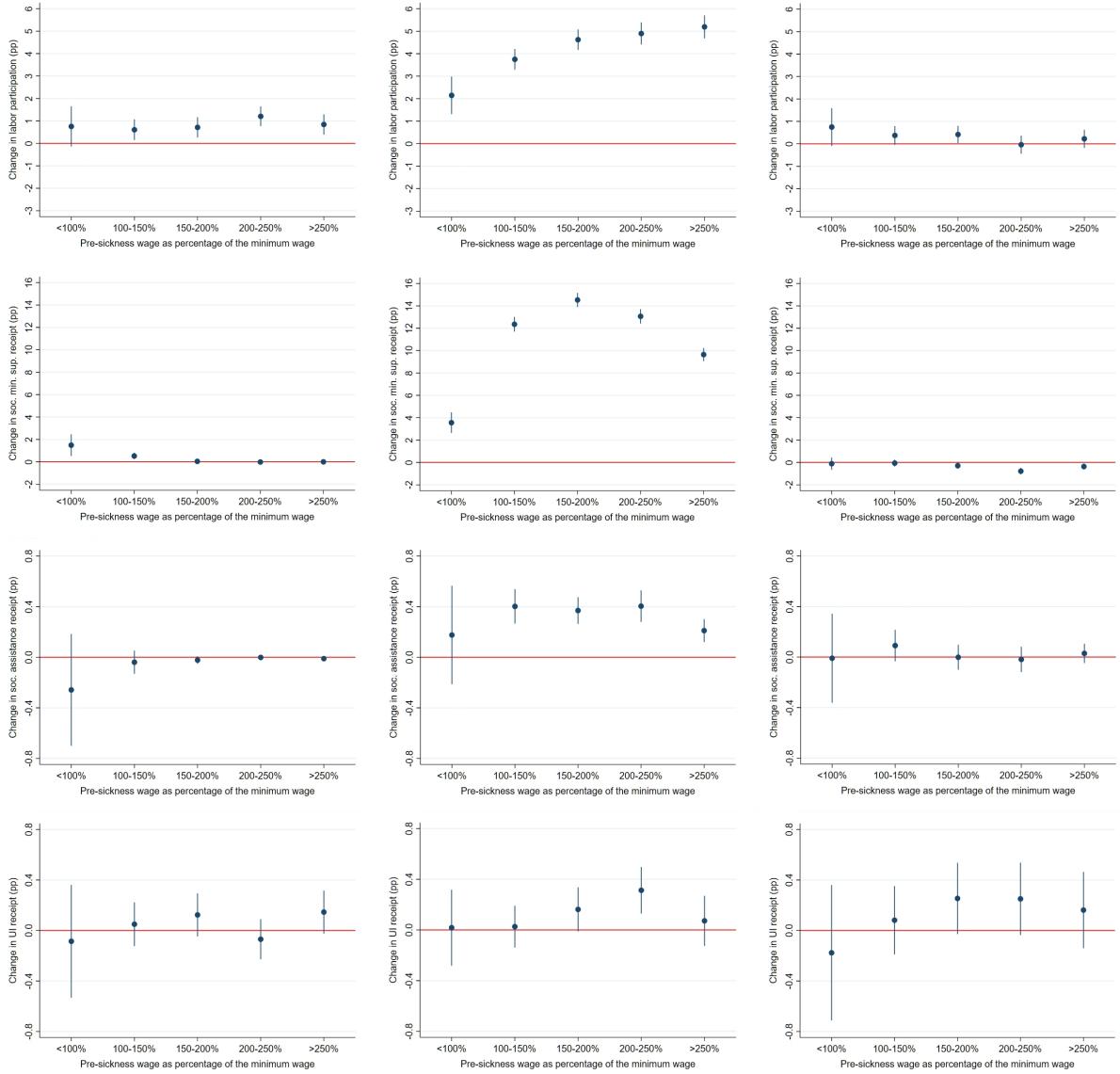


Figure A22: Estimated effects of the financial incentives on labor participation and receipt of three types of benefits and 95% confidence intervals around them based on the true cut-off (center panel) and placebo cut-offs 18 months before (left panel) and after (right panel) the true cut-off.

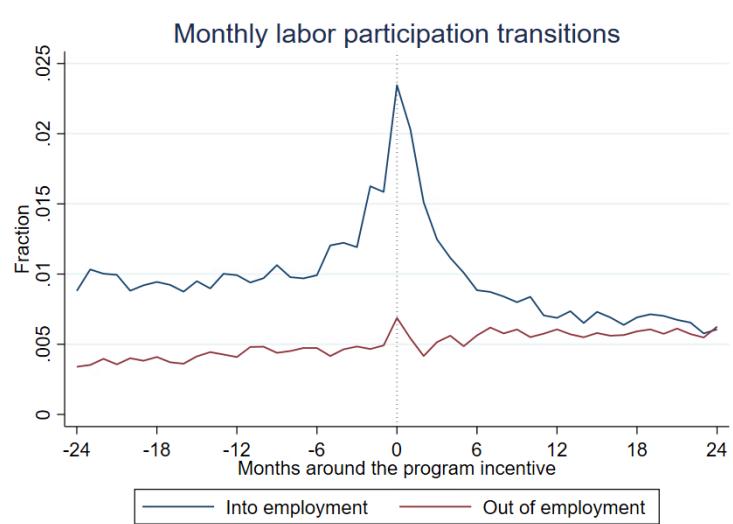


Figure A23: Fraction of people working (not working) conditional on not working (working) in the previous month during the last 24 months of the wage-related benefit and the first 24 months of the wage-supplement or the follow-up benefit when the financial incentives are effective.

## Covariate balance

If the local randomization assumptions hold, all factors driving the outcome variable other than the treatment indicator should not discontinuously change at the cut-off. Although this assumption cannot be tested directly, relevant covariates can be checked for whether they change significantly at the cut-off. We consider health as the most relevant covariate. In the Netherlands, the mandatory health care insurance (Zorgverzekeringswet) reimburses costs of medicines prescribed in a year. Information on prescribed medicines are available from Statistics Netherlands. Based on this information, we calculate the number of medicines prescribed in a year, and consider larger number of medicines prescribed in a year as an indicator of deteriorating health. We then estimate equation (1) using this variable as the outcome. A significant treatment indicator suggests that treatment assignment is not random.

Since the information on medicines prescribed is available on a yearly basis, it is difficult to test the impact of the treatment indicator based on (event) months, in particular at the cut-off month. That is, we need that values of the outcome can change at the cut-off month so that we can test whether the outcome exhibits a discontinuous jump at the cut-off month. To address this, we restrict the sample to individuals for who medicines are prescribed in different years before and at the cut-off month or later. For the smallest window around the cut-off, for example, this implies considering individuals who are prescribed medicines in December of a year in the last month before the cut-off (event month -1) and in January of next year at the cut-off month (event month 0).

Figure A24 presents the estimated treatment effects. Wider confidence intervals of estimated effects of wider windows are due to smaller number of individuals having observations at those windows. Number of medicines prescribed does not exhibit a discontinuity at the cut-off in any pre-sickness wage group or window except in rare cases. We note that, in our setting, it should be unlikely that covariates are imbalanced due to treatment since individuals face the treatment at random times depending on the number of years they worked before falling sick. Hence, any observed imbalance is likely be due to factors other than treatment.

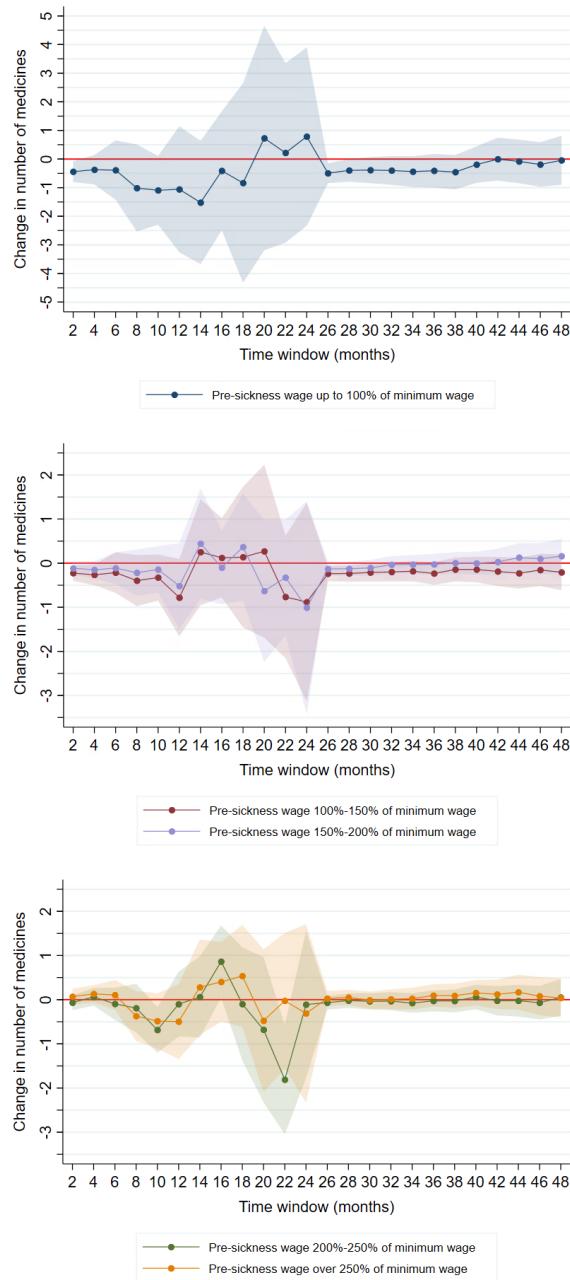


Figure A24: Estimated effects of financial incentives on number of medicines prescribed and 95% confidence intervals around them.

## Continuity-based approach to RD

We consider the widely implemented continuity-based (CB) approach to RD. Following standard practice, we consider a linear fit of the running variable, and a triangular kernel for weighting the observations centered around the cut-off. We choose the bandwidths to minimize the mean squared error (MSE) of the RD estimator. The confidence intervals for coefficient estimates use standard errors that account for heteroskedasticity and clustering at the individual level in all regressions. Regressions include the same covariates as in the baseline regressions based on the LR approach to RD. We also consider individual fixed effects as in baseline analysis. This means that our regression specification changes only in the approach to RD we take, that is, from LR to CB. The top-left panel of Figure A25 presents the estimation results. The results provide evidence of an increasing effect of the financial incentives across groups earning higher wages before falling sick, and they are consistent in the order of magnitude with the results from windows of 2 months based on the LR approach to RD in Figure A5. The top-right and the bottom panels of Figure A25 presents the estimated effects for benefit receipt. Again, the effects are consistent in the order of magnitude with those based on the LR approach to RD in Figure A6.

As in Section 5.2, assuming that anticipation and adaptation are part of the true responses, we investigate to which extent they lead to an underestimation of the true effect of the financial incentives in the baseline regression analysis (based on the CB approach to RD). As in Section 5.2, we consider donut hole regressions where we exclude observations within the window of 8 months around the cut-off month, keeping otherwise the baseline regression specification the same. Figure A26 presents the results. The treatment effect estimates are substantially larger than those based on the baseline regression (of CB approach to RD). This suggests that the baseline treatment effects are underestimated due to anticipation and adaptation effects. Compared to the LR approach to RD in Figure 8, the effects remain substantial although they are about 2 pp smaller for labor participation and social minimum supplement receipt.

We repeat the checks on identifying assumptions conducted for the LR approach to RD in Section 6.3. Figure A27 presents the estimated effects of financial incentives at placebo cut-offs, 18 months before and after the true cut-off, keeping otherwise the continuity-based RD specification the same as in the baseline. None of the placebo effects are significant, providing supporting evidence that the significant effects observed at the true cut-off are due to the financial incentives and not other factors.

If the continuity assumptions of the continuity-based approach to RD hold, the treatment should affect only the outcome at the cut-off and not any covariate. To test this assumption, relevant covariates can be checked for whether they change significantly at the cut-off. Considering health as the most relevant covariate, we repeat the continuity-based RD estimation replacing the outcome with the number of medicines prescribed in the year concerned. A significant treatment indicator suggests that treatment assignment is not random due to the outcome. Figure A28 shows that the treatment has no significant effect on medicine use for three highest wage groups but has a significant but small effect for the lower wage groups.

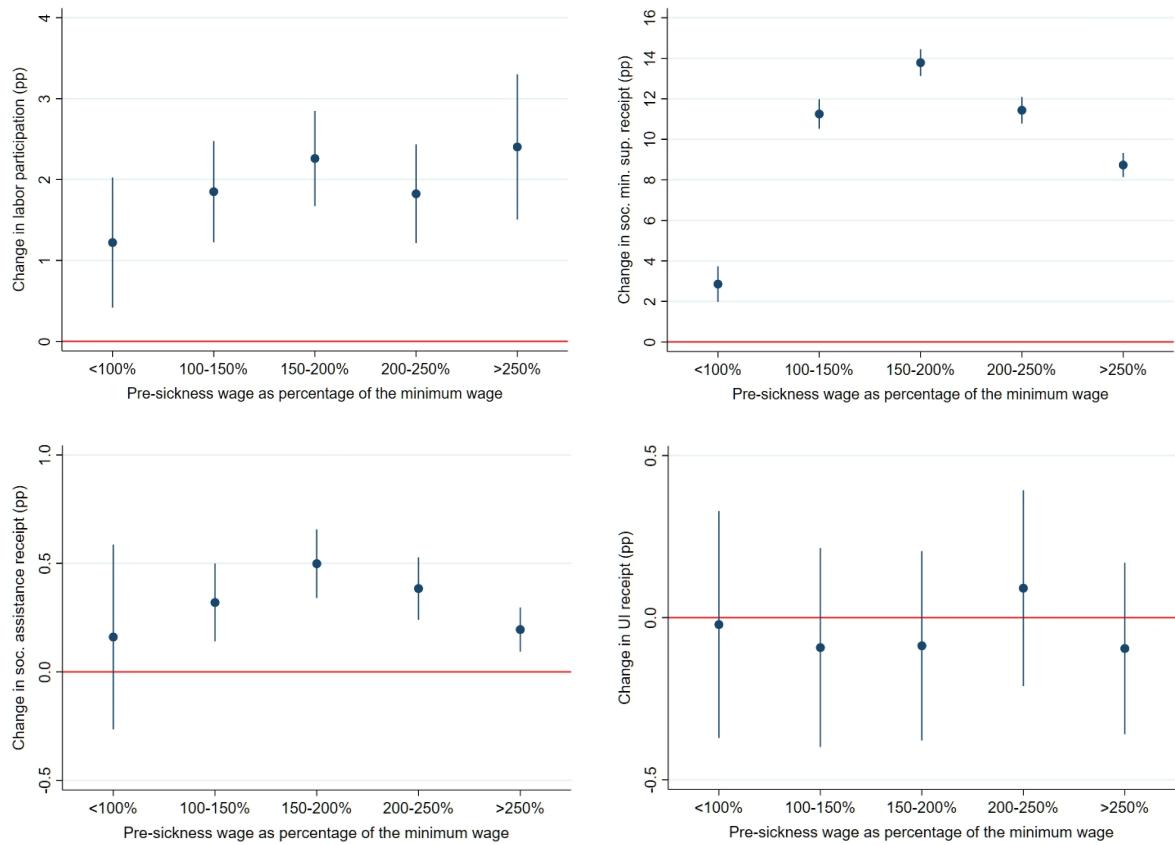


Figure A25: Estimated effects of the financial incentives on labor participation and receipt of three types of benefits and 95% confidence intervals around them based on continuity-based regression discontinuity design.

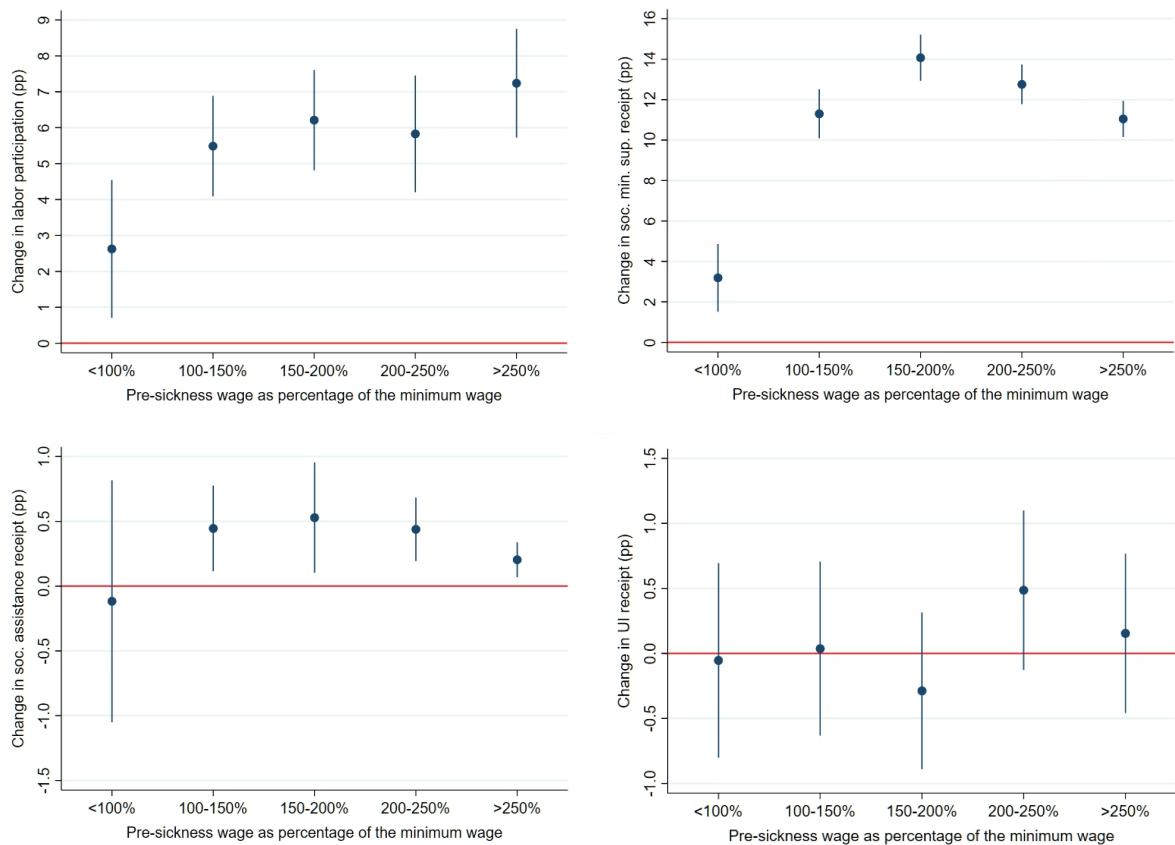


Figure A26: Estimated effects of financial incentives and 95% confidence intervals around them based on the donut hole regressions.

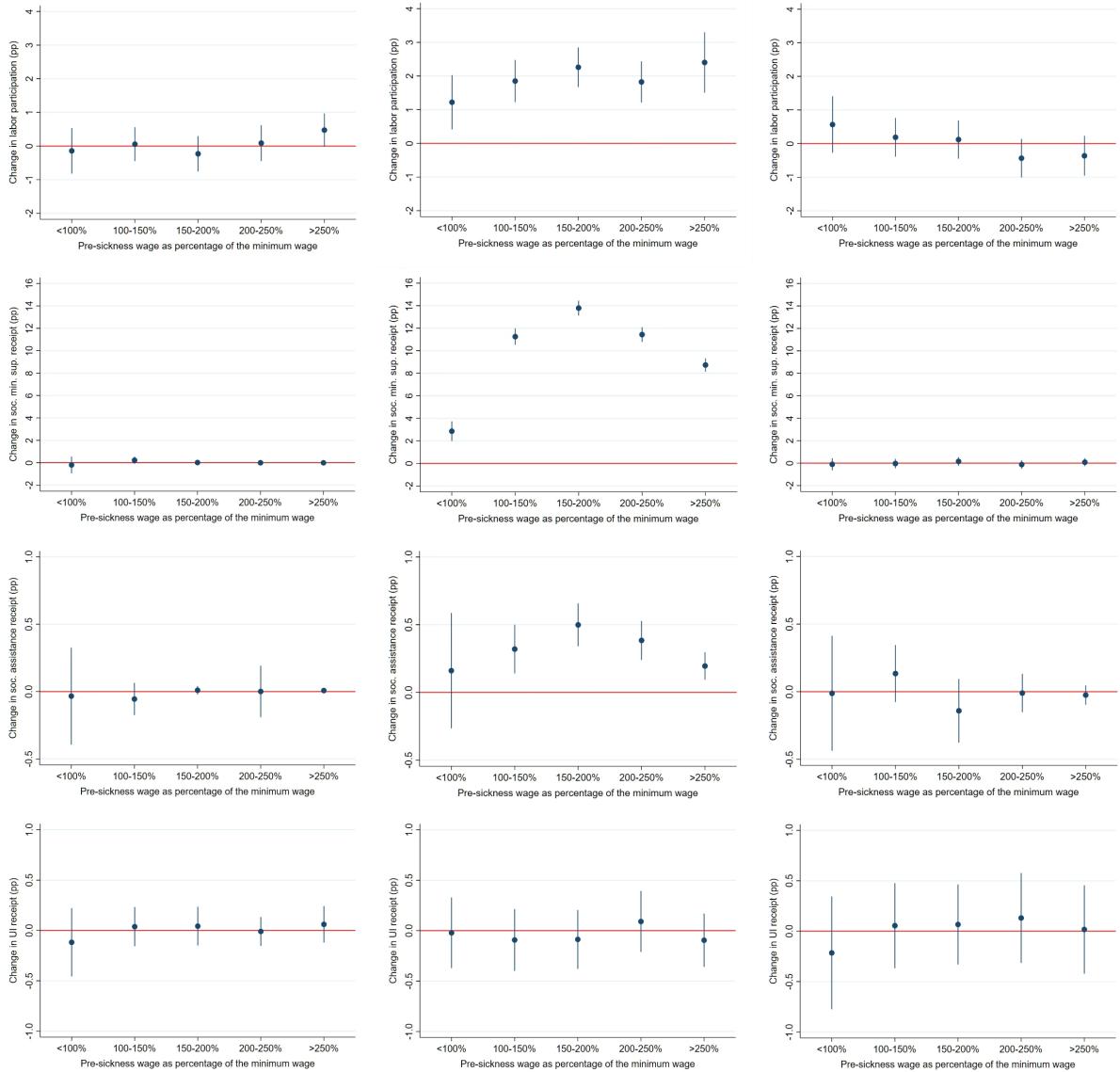


Figure A27: Estimated effects of the financial incentives on labor participation and receipt of three types of benefits and 95% confidence intervals around them based on the true cut-off (center panel) and placebo cut-offs 18 months before (left panel) and after (right panel) the true cut-off.

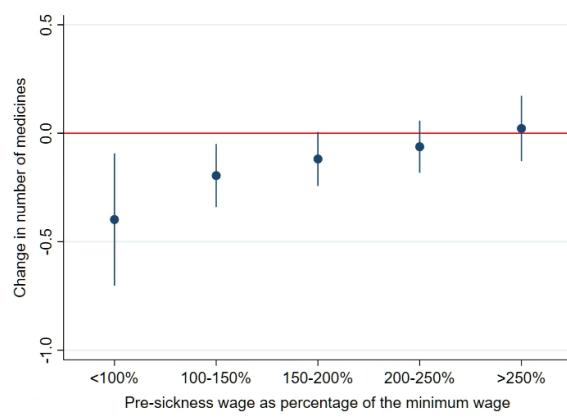


Figure A28: Estimated effects of financial incentives on number of medicines prescribed and 95% confidence intervals around them.

## An alternative identification strategy

To analyze the causal impact of the financial incentives of the WGA, [Koning and van Sonsbeek \(2017\)](#) exploit the variation in the timing of facing the financial incentives at the cut-off, and compare the labor participation of “treated” individuals who are close to or after they face the financial incentives with those of the “control” group of individuals who are not or not yet affected by the financial incentives. This identification strategy relies on two assumptions. First, treatment and control groups share the same time trend in the potential outcome throughout DI claiming. Second, relevant behavioral changes occur in a specific window around the month financial incentives take effect. This implies that DI recipients are able to respond to the financial incentives at reasonably short notice.

The regression model includes two dummies that capture the short- and long-term effects of the financial incentives. One dummy indicates a window of 6 months (in the benchmark model) around the month financial incentives take effect, and another dummy indicates the succeeding months. Other preceding months are chosen as base. Coefficients on these dummies represent the short- and long-run treatment effects, respectively. Other controls include five-year age group dummies interacted with a fourth order polynomial of the time spent claiming DI, calendar year dummies, and individual fixed effects.

We fit this regression model on our data, and compare the resulting treatment effect estimates with those based on the RD model we specified in equation (1). Figure A29 presents the estimated short- and long-term effects of the financial incentives. These estimates are close to our estimates from short and wide windows in Figure A5.

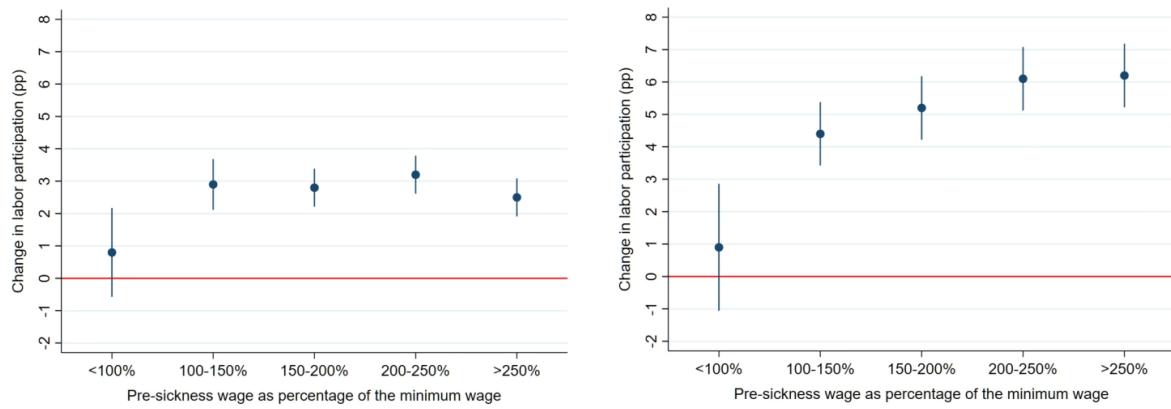


Figure A29: Estimated effects of the financial incentives on labor participation in the short (left panel) and long (right panel) run and 95% confidence intervals around them. The identification strategy is based on that of [Koning and van Sonsbeek \(2017\)](#).